TAB BOOKS / No. 583

$5.95

INDUSTRIAL ELECTRONICS PRINCIPLES & PRACTICES

By Alfred Haas

INDUSTRIAL ELECTRONICS Principles & Practice

By Alfred Haas

TAB BOOKS
Blue Ridge Summit, Pa. 17214

FIRST EDITION

FIRST PRINTING—DECEMBER 1971

Copyright © 1971 by TAB BOOKS

Printed in the United States
of America

Reproduction or publication of the content in any manner, without express permission of the publisher, is prohibited. No liability is assumed with respect to the use of the information herein.

International Standard Book No. 0-8306-1583-X

Library of Congress Card Number: 78-178690

Preface

Electronic—the magic of the age we live in—is a rather sophisticated, continuously expanding domain penetrating an increasing number of human activities: entertainment, communications, defense, scientific research, medical diagnosis and therapeutics, banking and insurance (computers) and, last but not least, industry. In but two decades, the transistor induced electronic technology and applications to proceed at far greater strides than did the vacuum tube in more than half a century; for, while tube techniques were adequate for putting together receiver circuits comprising half a dozen tubes, only the transistor made practicable computers and other complex gear containing hundreds, and even thousands of discrete devices. And this trend is progressing even faster now, due to the development of integrated circuits, designed for assembling compact tabletop computers, as well as large systems to control industrial plants, space vehicles and other worlds of complexity.

The growing number of rapidly maturing specialized electronic fields adds up to a more and more confusing pattern. Many excellent textbooks and technical papers devoted to specific matters of limited scope are available, and you can't do without them if you wish to dig deeper into a given field. To paraphrase an old proverb, the trees can prevent you from seeing the forest, it is becoming increasingly difficult to get a bird's-eye view of the whole electronic domain.

This book is an attempt to present an overall picture of a rather broad field: **Industrial Electronics**, the bulk of electronic engineering as applied to industrial operations. However, if a bird's-eye view embraces an extended landscape, it lacks detail; it would take a large series of books to exhaust so broad a subject. So, it is clearly stated that no claim as to exhaustiveness, or nearly so, is made. No attempt is made, either, to provide "ready-mix" prescriptions for solving any possible problem, because such problems are countless. Rather, we endeavor to sift out the fundamental principles of industrial electronics and to show approaches to

problems encountered. This is an essentially practical approach, emphasizing the understanding of underlying physical realities rather than purely theoretical material and involved calculus.

The Basic Industrial Electronics Course by the author, published by Gernsback Library in 1962, is the forerunner of the present book. However, while the former book mainly dealt with tube circuitry, the switch to semiconductor techniques is achieved in the present work. Also included are recent technologies such as integrated circuits and fields such as opto-electronics. Therefore, this is not just a new revised edition of the 1962 book, but a completely rewritten and largely augmented work. Nearly all material contained is new.

Alfred Haas

Contents

1 INDUSTRIAL ELECTRONICS: PURPOSE & MEANS 7

What is Electronics?—Electronics, A Tool for Modern Industry—Reliability—Open- and Closed-Loop Systems—Regulators and Servosystems—Transient Behavior of Servomechanism—Optimizing Control-System Operation—Null Methods for Accurate Measurement of Process Variables—Potentiometer Method for Voltage Measurement

2 TRANSDUCERS 39

Variable Resistance Position Transducer—Strain Gauges—Resistive Temperature Transducer—Liquid Conductivity and pH Measurements—Capacitive Transducers — Inductive Transducers — Angular Position Transmitting Synchros—Pulses for Revolution Counting — Tachometer Generators — Photoelectric Transducers—Opto-Electronics—Photoelectric Controls—Opto-Electronic Link or Modulator—Hall-Effect Transducers.

3 SYSTEM BUILDING BLOCKS 105

Amplifier Types and Characteristics—Inverse Feedback—Vacuum Tube Amplifiers—Direct-Coupled Electron Tube Amplifiers—The Transistor—Transistor Amplifiers—Thermal Compensation of Direct-Coupled Transistor Amplifiers—Field-Effect Transistors (FET)—Microelectronics—Carrier-Current Amplifiers—Chopper Modulators—Switching Behavior of the Transistor—Impedance Converters—Feedback Oscillators and Wave Generators—Multivibrators and Related Circuits—Other Important Circuits—Phasing Circuits—Transformation of Waveforms—Hot-Cathode Thyratrons—Cold-Cathode Thyratrons—Thyristors and Other Solid-State Switching Devices—The Unijunction Transistor (UJT)—Nonlinear Devices

4 AUTOMATIC INSPECTION, SORTING & COUNTING 184

Structure of Inspection Systems—Stepping Switches for Programmed Inspection—Automatic Inspection of

Multiconductor Cables—Automatic Testing of Passive Components—Testing Chassis and Subassemblies—Self-Balancing Bridges and Potentiometers—Galvanometric Amplifiers—Automatic Dimensional Inspection—Automatic Quality Control—Electromechanical Counters—Electronic Counting and Numeral Display Devices—In-Line Numeral Displays—Preset Electronic Counters and Digital Time-Delay Generators—Events-Per-Unit-Time Counter—Digital Voltmeters—Binary Counting—Eccles-Jordan Counter Decade—Binary-Coded Decimal Counter—Production Line Counting

5 DIGITAL PROCESS CONTROL 234

Logic Operations—Tracing Logic Diagrams—Static (Solid-State) Switching—Digital Integrated Circuits—Decoders for Display Driving—Static Controls Using Magnetic Devices—Analog-to-Digital Conversion—Program Control—Numerical Position Controls for Machine Tools—Adaptive Controls—Analog Computers—Automatic Control of Complex Processes or Automation

6 ELECTRONIC HEATING, WELDING & MACHINING 294

High-Frequency Heating—Induction Heating—Electron-Beam Heating—Electronic Resistance-Welding Controls—Synchronous Low-Power Welding Control—Ultrasonic Processing and Machining—Ultrasonic Cleaning and Liquid Processing—Ultrasonic Machining—Ultrasonic Soldering and Welding—Spark Machining—Evaporation, Drilling and Welding by Basic Beams

7 ELECTRONIC SAFETY SERVICES 326

Press and Shear Safety Devices—Making Safety Devices Safe—Flame-Failure Protection—Fire and Smoke Detectors—Intruder Alarm Systems—Proximity Switch for Preventing Collisions—Gamma and X-Ray Radiation Monitoring

8 POWER CONVERSION & CONTROL 343

Rectifier Types and Connections—Controlled Rectifiers—Regulated Power Supplies—Constant-Voltage and Constant-Current Supplies—Line-Voltage Regulation—Ignitrons—Saturable Reactors—Magnetic Amplifiers—IC Motor Control—Accurate DC Motor Speed Control—Squirrel-Cage Two-Phase Servomotor Control—Ward Leonard Drives—Amplidyne Rotary Amplifier

CHAPTER 1

Industrial Electronics: Purpose & Means

Before delimiting the scope of what is termed Industrial Electronics, we have first to answer the very fundamental question:

WHAT IS ELECTRONICS?

Although named much later, electronics was born with the invention of the vacuum tube as an offspring of the tree, then amply centenarian, representing the branch of physics called electricity. Electronics obviously got its name because electrons are involved; but since the electric current passing in electric motors and other long known devices is made up of electrons, too, we will try to draw a frontier separating the two domains. As electrons invariably are involved, the distinction is bared on the medium they pass through. In electric circuits, this medium nearly always is a metal, mostly a good conductor such as copper. During the vacuum tube area, the medium characterizing electronics was the vacuum or gas that supported conduction in the bulb.

With the invention of the transistor, semiconductor materials have become a typical electronic medium. But now, what about liquids? Fairly conductive solutions called electrolytes have long been used for metal plating and admittedly belong to electricity. On the other hand, oldtimers may remember the electrolytic rectifier that preceded the thermionic and solid-state devices, and a comeback can never be foreclosed. Avoiding too definite a statement, let us tentatively put it this way:

"Electronics is the science and technology dealing with the movement of electrons in mediums other than good electrical conductors."

Coming down to earth, a circuit is termed electronic if it contains electron tubes and-or semiconductor devices such as transistors, diodes and integrated circuits. Most of these devices are active; i.e., capable of amplifying or, generally

speaking, transforming electric signals. Owing to the particular ease of signal amplification and transformation by compact and economic means, electronic techniques are outstandingly flexible. But electronics always remains a branch of the electricity tree rather than an entirely different domain. And this is especially true in industrial electronics where an electronically controlled current often powers devices belonging to conventional electricity such as motors and relays.

ELECTRONICS, A TOOL FOR MODERN INDUSTRY

The tremendous development of industry certainly is the most remarkable feature of our era. Progress in transportation and communications, among other economic and social factors, has pushed industry into mass-production schedules to serve expanding markets, and goods must be produced economically to meet competition and customer spending willingness. On the other hand, high living standards have made labor an expensive item not to be wasted, not to speak of the shortage of qualified men. For all these reasons, industry is continuously looking for new technologies to produce more, better and cheaper.

In solving production problems, electronics can aid industry in various areas. Electronic devices provide for precise dimensional stability of mass-produced workpieces, they inspect and sort, gauge and count components and products, and they protect operators and facilities. All these applications can conveniently be considered as **electronic aides to manufacturing.**

Going a step further, the transfer of the work from one machining station to the next can be made wholly automatic, and still further, all the transfer mechanisms and individually controlled machines can be integrated into a unique computer-controlled process (or production) control. This is called **automation.** In such an automated system, production is carried out without any human intervention, or nearly so. (From another point of view, the malfunctions of any one element of the loop can lead to a breakdown of the whole production line. Skilled technicians, therefore, must be on hand to quickly put things right so that production can be resumed without costly delays. These experts will continuously provide preventive maintenance to the whole system to avoid such breakdowns.) Thus, automation can be termed **electronic production control.**

In addition, a lot of processes such as heating, fusing and welding, cutting and drilling, precipitation of paint, vacuum

deposition and others are frequently carried out by electronic methods. Though expensive, electronic processing can be more convenient, and thus even more economical than conventional methods, and sometimes this is the only way practicable. Production of very delicate (and costly) items such as semiconductor devices often requires electronic equipment performing in very specially controlled ambient conditions where conventional tools are useless. Let us term this class of applications as **electronic tools.**

Notice that the electronic solution of control problems often is not the only one practicable; there may be purely electrical, mechanical, pneumatic and hydraulic ways to solve them, and they possibly may result in simpler, cheaper and more reliable equipment. Thus, for instance, pneumatic and hydraulic devices are often preferred in explosive ambients for safety reasons; fluid controls are a natural for this application. But non-electronic controls are mostly bulkier and costlier than their electronic counterparts, and this particularily applies to remote controls, for the electric signal is the easiest to transmit.

RELIABILITY

If your domestic TV set goes silent, it is generally no matter for deep concern. If you have a second set, you will not even be deprived of your favorite program. On the other hand, failure of an industrial electronic control can stop a production line and, therefore, be very costly. That is why production managers and foremen considering setting up an electronic control ask awkward questions about reliability.

Such distrust once was justified, for early industrial electronic equipment was designed much like a radio set, using the same type of poorly engineered components. As a result, breakdowns frequently occurred. Recognizing the seriousness of this shortcoming, industry has since conducted comprehensive reliability programs, and, owing to this tremendous effort, components, materials and equipment have been upgraded to a point where failure is practically eliminated. As a fallout of space technologies, parts and systems of utmost reliability have been marketed. Over-critical users should remember that the human operator replaced by an electronic control does not have an unlimited reliability record either, and, as a general rule, the electronic system is more trustworthy.

There are no special reliability techniques, but rather applications of **good engineering practice.** For one, overrating considerably shortens the service life of a component. For

instance, a 1-watt resistor is capable of dissipating 2 watts or more for some time, but inevitably the resistor will deteriorate more quickly. On the other hand, underrating (i.e., using a 2-watt resistor to dissipate 1 watt) results in an almost infinite service life. Notice too, that the rating specified by the manufacturer holds only for a given ambient allowing for adequate cooling. A resistor working in a high temperature ambient or deprived of cooling air draught must be derated to avoid rapid deterioration. Forced-air ventilation can enhance reliability.

This emphasizes the necessity of defining the limits of the peculiar operating conditions that may vary widely from case to case. Obviously, it is unnecessary and unduly costly to design for operation in arctic cold an equipment intended for use in an average factory; but on the other hand, if the device is to be bolted to a machine subject to vibrations, the amplitude and frequency characteristics of the vibrations must be known to permit designing the equipment accordingly.

There are special "reliable" tubes of sturdy design built to resist high shock and vibration levels, and other types that feature very uniform characteristics and long life expectancy. The particularily careful processing and checking required to produce such tubes add up to make them more expensive, but the higher probability of long, trouble-free operation is worth more than the added expense. Diodes, transistors and other solid-state devices are inherently more reliable, since there are no heaters to burn out and the compact and sturdy structure makes them more immune to shock and vibration than vacuum tubes. Also, the negligible heat dissipation with transistors makes for cooler and thus more reliable operation. Semiconductor devices, however, are serviceable only at moderate temperatures, while tubes that will endure high temperatures are available. Furthermore, a transistor is easily ruined by a short current or voltage transient that a tube will painlessly "swallow." Thus, there still are applications where the tube has an edge on the transistor.

Most passive components have reliable counterparts. There is a lot of difference, for example, between a common "pot" for a TV set and a wirewound precision potentiometer for analog computers. Potting and impregnating confers immunity against humidity, corrosive atmosphere and biological damage (by fungi). Good solder joints are of utmost importance to reliability. Even if "cold" joints obviously are corrected, a good soldered connection can corrode if the solder used was cored with a too activated (acid) rosin to facilitate soldering. For that reason, subassemblies for critical applications are often welded rather than soft-soldered. In-

tegrated circuits with built-in connections drastically reduce the number of joints to be made on assembly and, therefore, increase reliability. (The inner connections are, of course, carefully tested before assembly.)

Mean Time Between Failures Specifications

As unlimited reliability factors conceivably cannot exist, manufacturers specify a given equipment a mean time between failures (MTBF) of, say, 2,000 hours in a given set of ambient conditions as to temperature, humidity, shock, vibration and-or corrosive atmosphere. This does not mean that the equipment will break down in the 2,001st hour; in fact, it may operate for 10,000 hours without failure. But reliability research has shown that faultless operation cannot be granted for more than 2,000 hours. (An elapsed-time meter may total the effective operating time.)

At the end of the MTBF period, a maintenance team will thoroughly check all parts and components, and particularily those subjected to wear such as relays and switches. Tubes testing weak will be replaced. (In critical applications, all of the tubes will be replaced, whether they test weak or not.) If the maintenance operations have been carried out according to the manufacturer's prescriptions, a new reliability period of the same duration as the first one can be expected.

With the replacement of most electron tubes by solid-state devices and low-power relays by static controls, there are fewer and fewer parts remaining that are subject to mechanical wear. Therefore, the MTBF is drastically lengthened, and there are fewer moving parts to look at, which simplifies maintenance operations.

Redundancy

It has been said that underrating and ample dimensioning are basic reliability concepts. There is another one: Since each part is a possible cause of failure, the surest way to enhance reliability is to reduce the number of parts as much as possible. For this reason, and not for saving a few dollars, many industrial electronic equipments are brute AC-powered directly from the mains or by means of an isolation transformer, thus eliminating rectifiers, filter capacitors and an iron-cored choke. A well-filtered DC supply is a must in all consumer electronics sets, but a strong hum component may not be objectionable here.

Redundancy (duplication) is a typical reliability technique. In Fig. 1-1A, switch S closes the load circuit. Un-

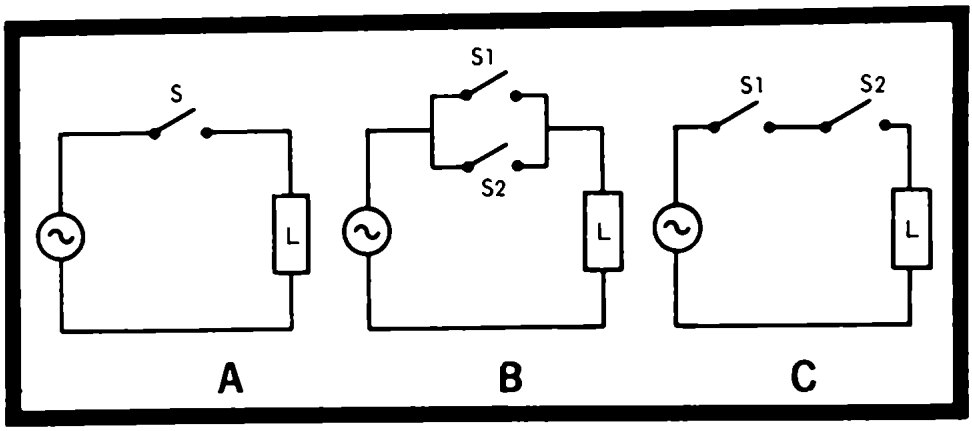


Fig. 1-1. Parallel switches (B) increase the reliability of a single switch (A), while series switches (C) operated simultaneously have a higher breaking reliability.

fortunately, switches often have a limited reliability record, and if S fails to close, load L is not energized. This may not be of utmost importance, and then, switches nearly always work. But if the switch closure is absolutely mandatory, be it for safety reasons, switch S1 is paralleled by another switch (S2) operated simultaneously (Fig. 1-1B). Since the result is now satisfactory if either S1 or S2 closes, reliability is two times greater now; in fact, it is much higher, for there is very little possibility for two rare events to coincide.

Another problem arises when the opening, and not the closure, of a switch is mandatory; for instance, consider an emergency interrupt. If the contacts stick together, the circuit will not be broken, and this opportunity is twice as great with paralleled switches. So we chose series connected switches (Fig. 1-1C); should one stick, the other almost certainly will open and cut the load circuit.

To avoid untimely shutdowns in broadcast transmitters and communication links, a spare transmitter is often operated in a standby condition, with heaters emergized, to take over immediately should the operating transmitter fail. If even the brief down-time necessary for automatic changeover cannot be tolerated, the tubes can be paralleled like the switches in Fig. 1-1B. Should one tube fail, the other would continue operating, and the transmitter would remain on the air, though at reduced power.

Such redundancy obviously is expensive, but in vital control circuitry where continued faultless operation is all-important, with cost being a secondary factor, it may actually be economical to provide redundancy for key components and subassemblies. Periodic checks must be made to make sure that both elements paralleled are in correct working order, for

failure of one of them may be hidden when the other takes over.

Ease of Servicing

Since an unlimited, failureless service life cannot be warranted, and because the fault has to be cured immediately to avoid lengthy work stoppages, industrial equipment is designed for easy access and-or replacement. Rack-mounted assemblies are often provided with elaborate sliding devices that allow a chassis to be pulled and rotated into a favorable position for inspection. Flexible supply cables of sufficient length allow tests to be made with applied operating voltages. Major components are (or should be) clearly labeled for easy reference to electric and topologic schematics contained in the service manual. A well-designed assembly will save time during routine maintenance operations and cut down production loss during emergency work stoppages.

With power devices excepted, modern solid-state electronics is compact and low heat-dissipating, allowing for high density circuit packing. Components are assembled on printed circuits or into potted units called modules. Even gear of moderate complexity may use many such subassemblies. For ease of assembly, checking and replacement, the printed cards, drawers and boxes are fitted with prongs for plug-in connection into wired chassis and frames. The circuitry of these plug-ins often is too crammed to be repaired by a field serviceman; fumbling around with these tiny devices would show up costly in terms of production line stoppage. So the defective item is discarded and simply replaced with a spare. Such compact and maintenance-free spares can be conveniently stocked for replacement.

In such a system, pilot lights are provided, preferably, for quick identification of the faulty circuit to be replaced. In complex systems, a large number of testing and fault-display devices is required for rapid servicing. But a multiplicity of alarm devices, plus all the machine and control equipment protection circuitry (against unpredictable power failure or cooling water stoppage, for instance), add up to crowd the schematic of an electronic control system. It is expedient to redraw such an illegible diagram, eliminating all protective circuits and alarm devices; then the schematic will stand out clearly.

Continuous and Discontinuous Controls

The discontinuous control operating in an "on-off" way is the simplest solution to a control problem, for only a single-

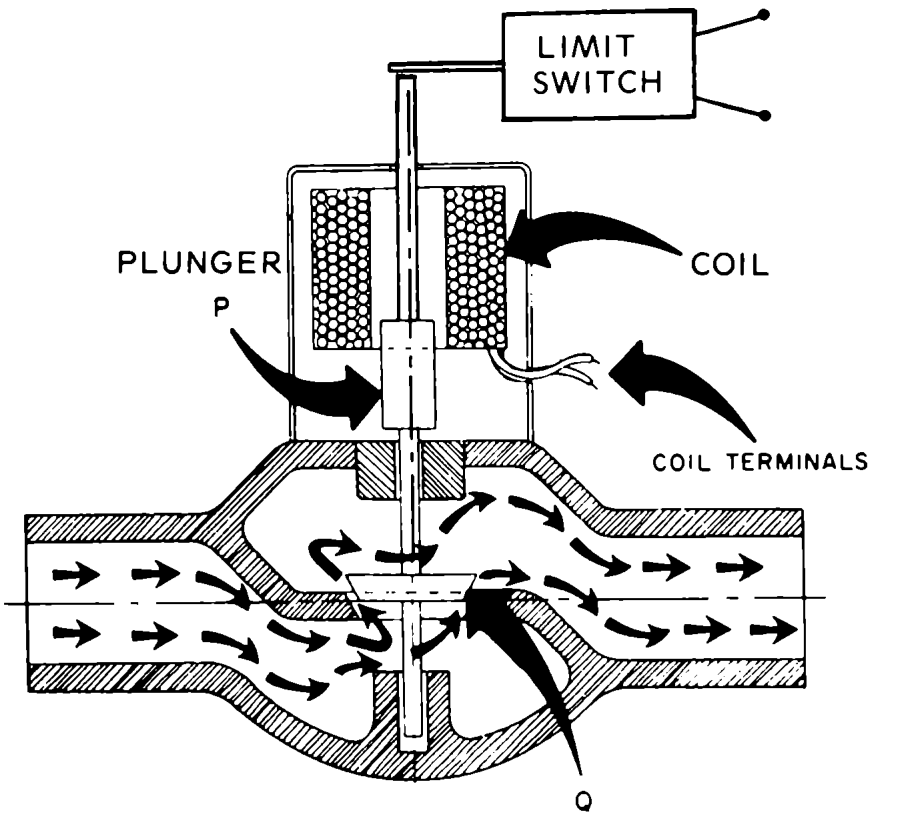


Fig. 1-2. A solenoid valve is a continuous on-off control.

pole single-throw (SPST) relay or switch is necessary to operate the device. The electromagnetic valve in Fig. 1-2 is a typical on-off control, for it can be only entirely open or completely closed. Plunger P is sucked in when coil C is energized, which lifts poppet Q and allows the working fluid to pass; interrupting the power supply results in the poppet falling back and a cessation of fluid flow. (There are many types of valves, and energizing the coil may act to start as well as to stop fluid flow.) Being fundamentally binary, as we will see later, this type of on-off control is a natural for digital systems and logic.

Such remote-controlled valves are extensively used in the chemical industries where they maintain a prescribed composition of a processed fluid. The use of a simple electromagnetic valve may, however, lead to prohibitive variations of the flow magnitude to be regulated. The motor driven valve in Fig. 1-3 allows for continuously variable flow control; between the limiting positions of "fully open" and "completely closed," there is an infinite number of intermediate positions. In this valve, the poppet is moved up and down by means of a spindle rotated by a geared-down motor.

(This scheme is used for demonstration purposes; commercial valves may be designed differently.)

At first glance, it is seen that such a continuous control is more complicated than the simple on-off device in Fig. 1-2. Means must be provided to have the motor rotate in the desired direction and to stop upon reaching a prescribed poppet position. The motor can be reversed by a double-pole double-throw (DPDT) relay which reverses the field of a series motor (Fig. 1-4); but static continuously variable reversing motor controls (to be examined later) are to be preferred. The poppet position can be sensed by a multiturn potentiometer (P in Fig. 1-3), driven by the screw spindle and connected in a null-seeking bridge which energizes the motor (through a power amplifier) until the balance point corresponding to the prescribed poppet position is reached.

Furthermore, such a motor-controlled valve has to be protected against overrunning the stops and thus damaging mechanical parts. This could be done with the potentiometer circuit outlined above, but a minor failure of the control circuit could make the protection inoperative. A good go-safe policy requires complete **independency of control and safety circuits.** Thus, the use of simple limit switches, acted upon by the axial spindle displacement, is preferred. As shown in Fig. 1-4, these limit switches (LS1 and LS2) are wired to shut off the motor if it attempts to rotate past the set limits, while still allowing for reverse rotation.

In some applications, a device ordered to carry out a given task is to report back to the order-issuing station if and how an order has been discharged; this is the basic philosophy of a closed-loop control. Such information can be provided by the potentiometer of the motor-driven valve. In the solenoid-operated device in Fig. 1-2, a relay series connected with coil C could sense the presence of driving current, thus indicating that the valve is opening. Imagine, however, that for some reason or other, the poppet mechanism has become sticky and remains open or closed. The relay current then no further reflects the actual condition of the valve. Therefore, if no false operating state indication can be tolerated, a limit switch LS can be installed on the solenoid to be acted upon by the spindle of the sucked-in plunger. Of course, there are very few chances for such a failure to develop in a well-designed valve; this example merely illustrates again the principle of separating control and safety circuits.

OPEN- AND CLOSED-LOOP SYSTEMS

Imagine an operator charged with filling a remote gasoline tank by operating a manual valve. This man can be

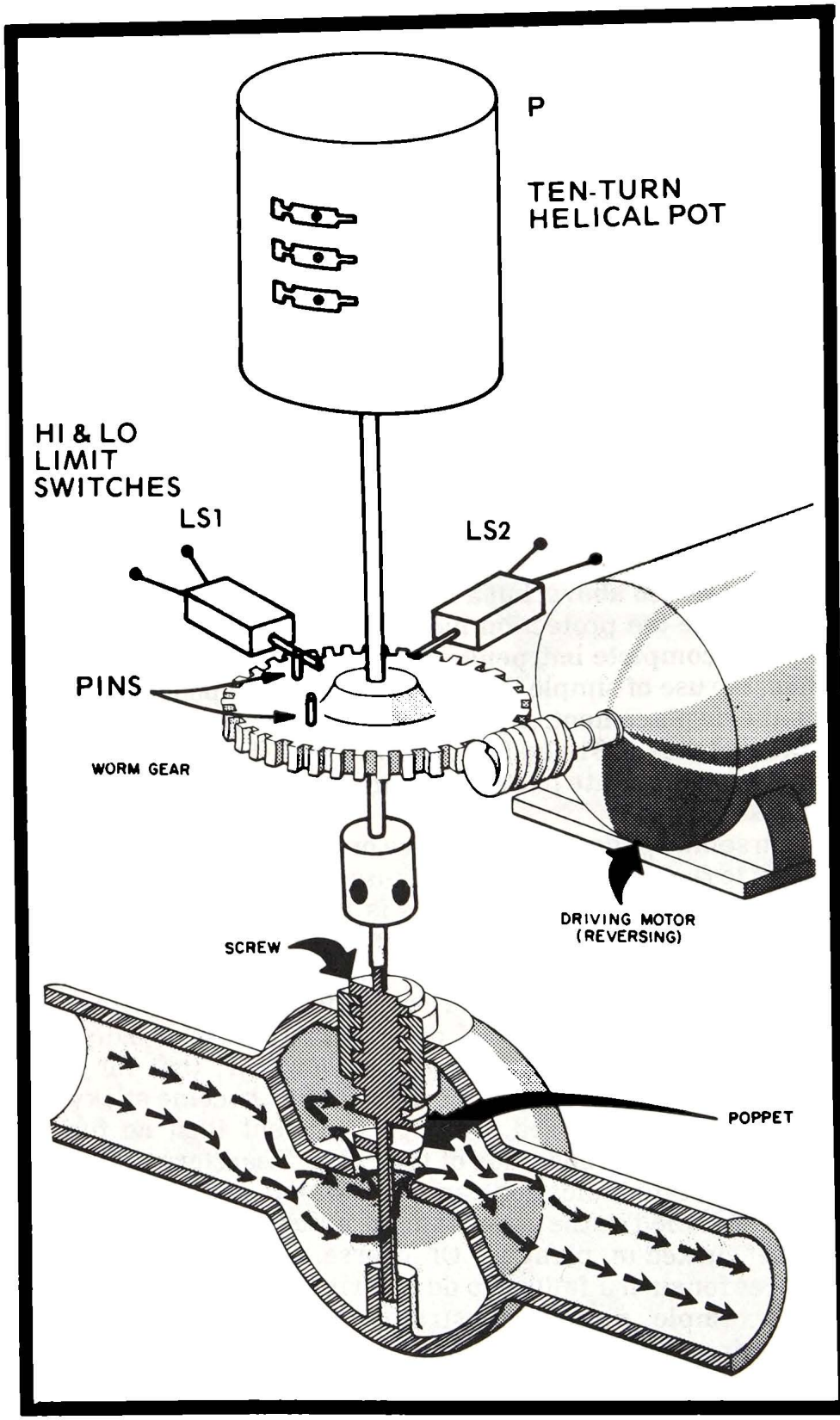


Fig. 1-3. A motor-driven valve provides for a continuously variable flow adjustment.

considered as an elementary valve **control system**. (Here is an example of poor utilization of such a complex and resourceful system as man; a motor or a solenoid could do his work.) If the operator is not provided a means for knowing what is happening inside the tank, he will be unable to stop filling when the gasoline reaches a specified level. This is an **open-loop** system, above all a theoretical case, for it cannot realize its objective. It is just like a general directing a battle who issues orders to his troops but is not kept informed of what is happening on the front.

To carry out his task successfully, the operator has to have a **means of knowing the result of his action** if he is to close the valve at the right moment. Just like the general, he could dispatch a messenger who can grasp the situation and report back to him. In control language, the messenger would serve as a **sensing device** or **transducer**. There, obviously, is another case of poor utilization of man power to report on such an elementary condition as a liquid level; inexpensive gauges would do the job better and cheaper. Getting this information closes the decision-making and action-taking loop necessary to successfully perform the proposed task. Now we have a **closed-loop system**. Considering technical as well as biological processes, you will be surprised to learn that the closed-loop principle governs most of them.

Considered as a control system, the human operator is unique in his ability to adapt himself to a given situation, relying upon previous experience, and to learn readily by doing. He will adjust his action to the desired result. Even a giant computer is not as intelligent, for it does only what its memory instructs it to do, and this memory merely contains what men have put into it.

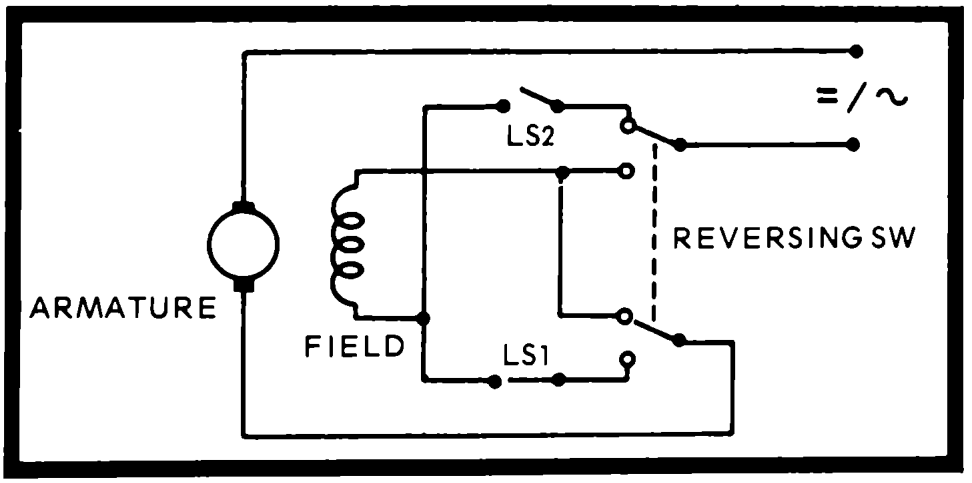


Fig. 1-4. A series DC motor circuit with reversing and limit switches.

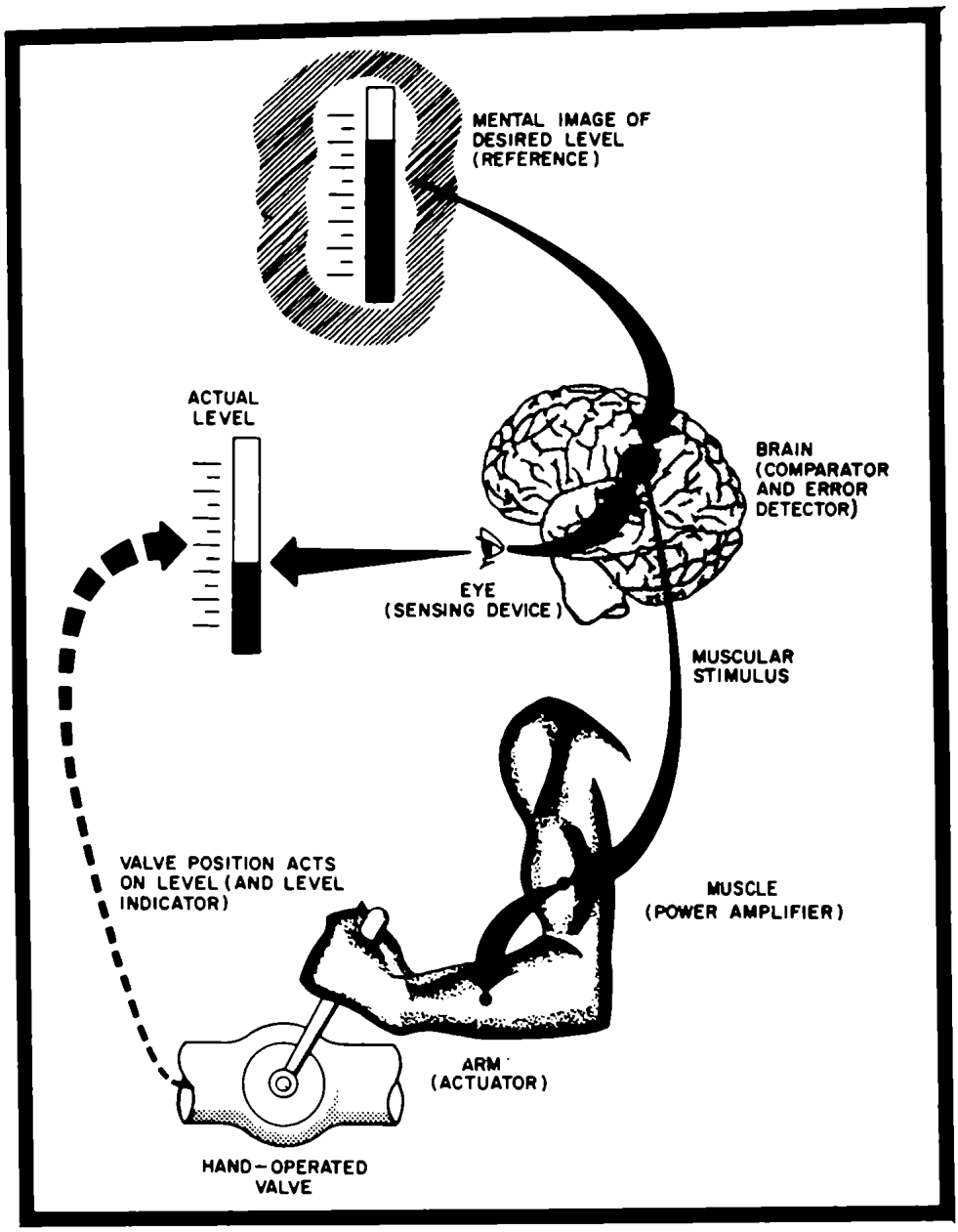


Fig. 1-5. A practical control system must duplicate these processes. In this drawing, a human operator is considered as a control system.

Let us now consider the above operator as a control system and study the coordination of his various functions (Fig. 1-5). Assuming that there is an optic level gauge, he will use his eyes as a **transducer** to evaluate the actual level. In his brain, this information is compared with the mental image of an indicator set to the desired level, or a **reference**. Acting as a **comparator** or **error detector**, the brain compares the actual and "ought-to-be" levels, decides on the necessary action and transmits an electrical **error signal** to the appropriate muscle.

The muscle, thus stimulated, acts as a **power amplifier**, and the arm becomes an **actuator**, opening or closing the valve. (The terms referring to control technology are printed in bold type to better stand out.)

Redrawing this pictorial scheme as a box-like diagram of the usual form (Fig. 1-6), we can represent all the essential elements of a closed-loop system as outlined. Letting out the sensing device (the reference and the comparator then become unnecessary) results in an open-loop system with limited use.

The closed-loop control system discussed is a very simple one. If you are looking for something more complicated, try analyzing the behavior of a human system as common as an automobile driver. Two major magnitudes must be controlled: speed and direction. But these depend on a number of variables that have to be sensed and related: road surface condition and congestion, visibility, climatic and ambient factors, behavior of other drivers, obstacles and crossings, and heaven knows what else, not to speak of all the unforeseeable conditions that suddenly may develop. The driver also has to consider the speed of his engine in relation to forward speed and choose the right gear (if there is manual gear shift); he has to sense the response of the car to turning, accelerating and breaking, not to forget his obligations for signaling and proper use of headlights. It would take a rather complicated computer to take into account all these variables for performing the actions of an average driver!

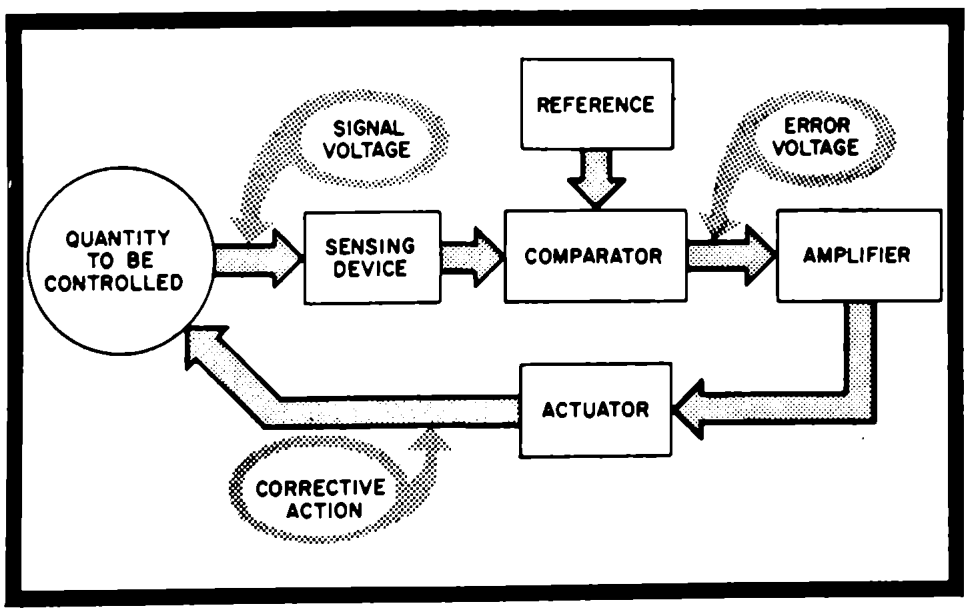


Fig. 1-6. This block diagram shows the relationship of the principal elements in a closed-loop control system.

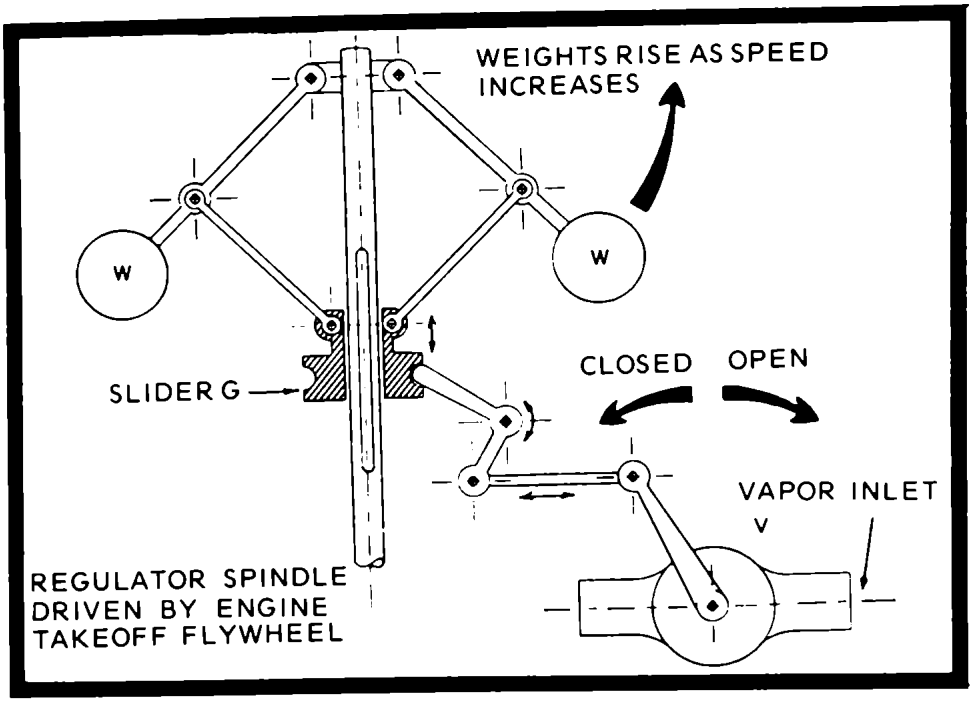


Fig. 1-7. Diagram of the Watt centrifugal regulator or governor—a mechanical closed-loop control system.

REGULATORS AND SERVOSYSTEMS

In all closed-loop systems, the magnitude to be controlled is compared with a reference quantity, and the control device then acts to minimize the resulting error signal. These two magnitudes may belong to different classes; for instance, a pressure (translated into a voltage) can be compared with a reference voltage. As in industrial electronics, we often change from one system to another; this makes no difference. A distinction has to be made, however, due to the fact that the reference can be a **constant** or a **variable** quantity. If the reference quantity is constant (or, at least, programmed), we are confronted with a **regulator**. This is the principal of James Watt's governor or centrifugal speed regulator for steam engines, invented some 200 years ago. In this purely mechanical control system, the spindle of the governor is driven by the engine flywheel. The resulting centrifugal force lifts the weights as the rotation speed increases (Fig. 1-7). A mechanical linkage transforms this motion into an asscending displacement of grooved slider G. This slider is linked to vapor inlet valve V in a way to throttle vapor as the weights ascend due to increased engine speed, and vice versa. This is a regulator, for its acts to maintain the engine at a preset speed.

The stabilized power supply in Fig. 1-8 is an electronic example of a regulator. Here, the reference magnitude is

voltage Vr stabilized by zener diode DZ in the emitter circuit of transistor T2. A voltage, V1, obtained from voltage divider R1P, supplied by the output voltage, is applied to the base of T2 and acts as a voltage comparator. T2 then issues an error signal and applies it to the base of series regulator transistor T1. T1 functions as a controlled variable resistor series connected with the load. Thus, if a load current increase tends to reduce the output voltage, and V1 by the same token (Vr remaining constant), T2 acts on base current of T1 to make it more conductive and to reduce dV. This increases the output voltage to restore its original level, preset by P. By applying a variable portion (V1) of the output voltage to the comparator, P allows for selecting the stabilized voltage level. If this level is fixed once for all, P can be replaced by a resistor.

In many automatic control systems, the magnitude under control is intended to follow a variable reference; we are now confronted with a **servomechanism**; i.e., a mechanism **slaved** to something. Such is the case of the copying lathe in Fig. 1-9. Work W rotates between the centers of the lathe as usual, and it is machined by a cutting tool mounted on a tool carriage. Positioning motor PM moves the tool back and forth towards W to modify the cutting depth; the actual diameter of the machined portion is sensed by mobile spindle transducer T1. A

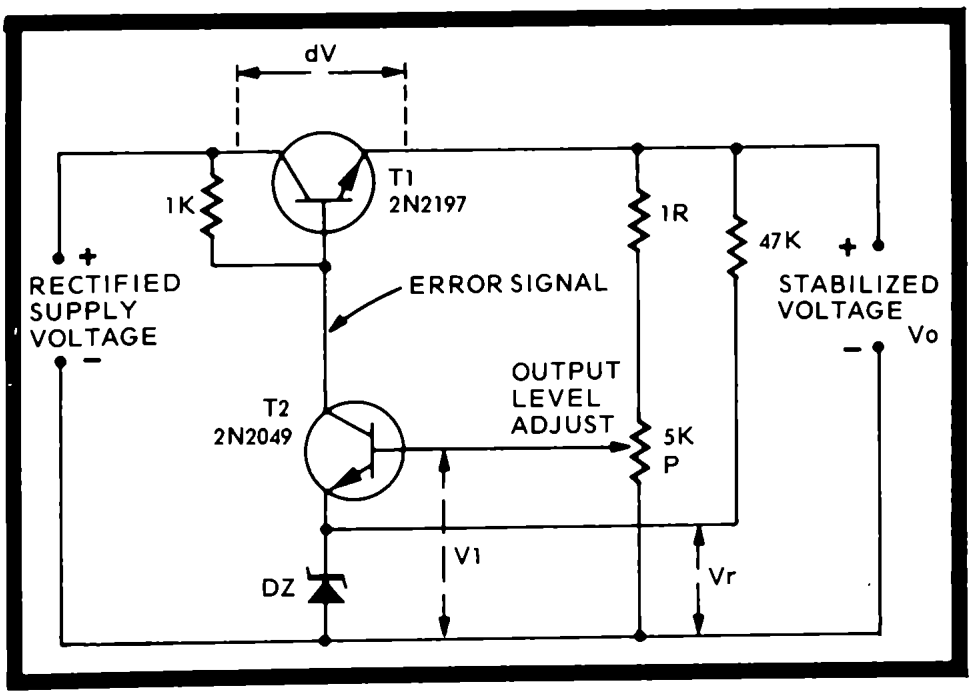


Fig. 1-8. This is an electronic regulator circuit designed to provide a stabilized voltage over a wide range of load drains.

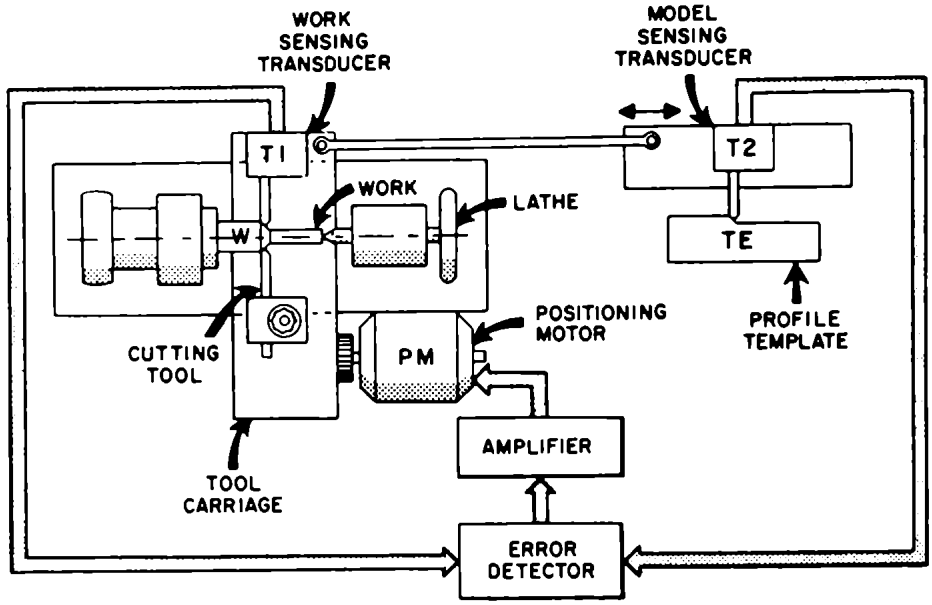


Fig. 1-9. A copying or duplicating lathe can be operated with the system diagrammed.

second longitudinally moving carriage, attached to the first one by means of a rigid linkage, senses the profile of stationary template TE by means of a second transducer T2. The signals of T1 and T2, representing the actual shape and the shape to be given to W, are fed to an error detector, and the resulting error signal is amplified to power motor PM to move the tool with regard to W. An unlimited number of identical items can thus be produced automatically, using a simple template. (The numerous auxiliary circuits and safety devices necessary for proper and reliable operation are not considered here.)

Instead of being generated again at each machining run by the template follower transducer, the variable reference signals produced by some type of transducer can be recorded on tape, together with the corresponding longitudinal position information, and used to replace the mechanical linkage. Such a **programmed control**, a step beyond the profile copying described, can be designed to allow for completely automatic production of large or small runs of identical parts. Supplementary signals can be recorded on the same tape to provide for tool and work change, positioning of work and other necessary functions. A "library" of recorded program tapes can be set up for economic machining of every part at least once fabricated, eliminating costly retooling. Only the very first run requires a skilled mechanic.

TRANSIENT BEHAVIOR OF SERVOMECHANISM

Consider a typical servomechanism such as the positioning control in Fig. 1-10. A sliding carriage (C) is displaced by motor M according to instructions issued by the control system. A transducer (T) senses the actual position of C and generates a position signal that is compared to a reference. The resulting error signal, once amplified, powers positioning motor M.

Let the carriage initially rest at position I. Suppose an incoming instruction orders the mechanism to move from I to final position II. If this order is issued by an on-off device such as a relay or a logic circuit, the corresponding command signal is a step function (Fig. 1-11A). This means that C, at rest at I up to instant t0, should instantly occupy position II and then be at rest again. No physical system is capable of such a performance. There are masses to be accelerated and inertia to be overcome, and acceleration cannot be infinite. On the other hand, a moving mass cannot be brought to a dead stop instantly. To the change proposed in Fig. 1-11A, the system, therefore, will respond by the variation shown in Fig. 1-11B. After a starting delay (t1 - t0), the motor will drive C to position II and reach it after a total time delay of (t2 - t0).

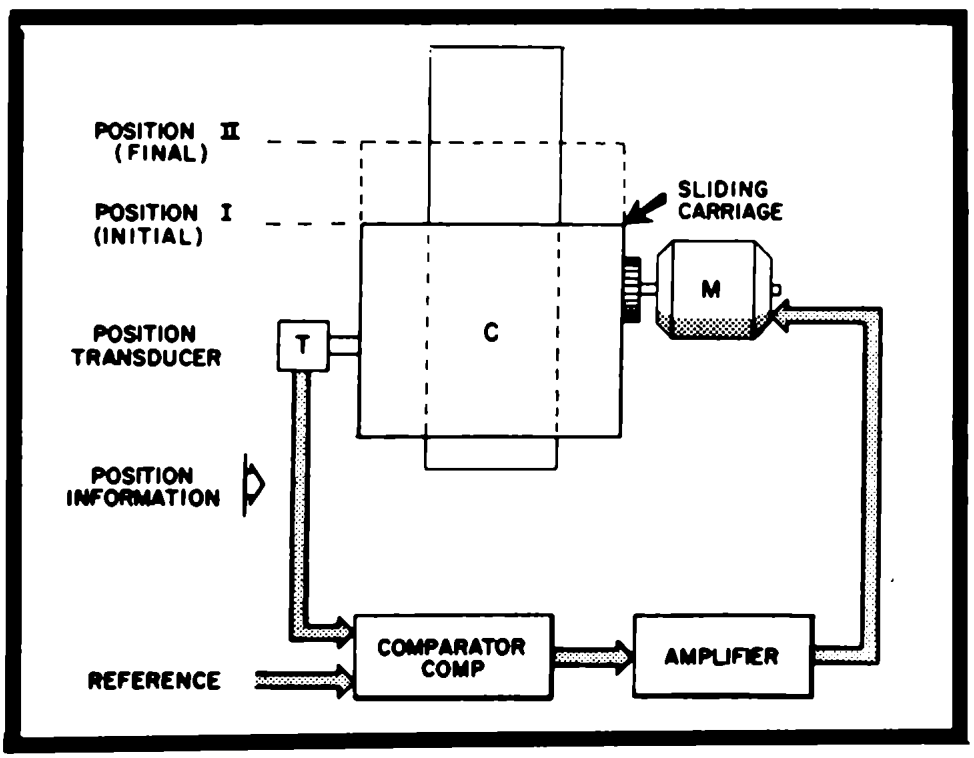


Fig. 1-10. Servomechanical systems deal with a time lag.

This characteristic describing the **transient behavior** (i.e., the response to a sudden change) is of utmost importance for the operating performance of the servomechanism, and hundreds of papers and books deal with its theoretical analysis and practical calculation. Such a time delay may be of no concern in a pseudo-static system such as a garage-door opener; it will, however, jeopardize correct operation of dynamic systems involving a rapidly changing reference quantity. Used as a tool-position control, the system can be ordered to drive the carriage from position I to II, and shortly afterwards, a new incoming order may specify a third position III (Fig. 1-12A). In this case, the carriage will reverse and drive towards III before ever having reached II. The actual displacement characteristic (Fig. 1-12B) differs greatly from the proposed function (Fig. 1-11A) leading, in this case, to considerable machining errors.

To make up for sluggish action, a more powerful motor, providing good acceleration and more effective braking, can be used, and the amplifier gain can be increased. This may result in a response to the unity step function as shown in Fig. 1-11C. Because of the rather steep characteristic, position II is now reached at instant t3, well before t2, but because of the inertia of the rapidly moving mass, the carriage "overshoots" its specified positon, stops, reverses, overshoots again, and so on. Unlike the former control behaving in a viscous (damped) manner, the present system "hunts" for its final position, tracing an oscillation of decreasing amplitude. Provided nothing interferes (such as a new order), the system will come to rest only at instant t4. These damped oscillations ("ringing") denotes the limit of stability. This limit is passed if the oscillation is of a constant or even a growing amplitude (Fig. 1-11D). The carriage then will keep on crossing the proposed position line indefinitely, being unable to stop at it and to come to rest. Such an oscillating system, obviously, is of no use.

Avoiding any overshoot leads to **over-critical damping** (Fig. 1-11E) with sluggish operation. A much quicker response without ringing is obtained by **critical damping** (Fig. 1-11F) described by a comparatively steep rise followed by a slight overshoot so small that the system then tends to reach the desired position without the slightest overshoot on the opposite side. This makes for stable operation, minimum control delay and accurate positioning.

OPTIMIZING CONTROL-SYSTEM OPERATION

To get the benefits of both a steeply rising control slope and slow variation near the final position to avoid overshoot,

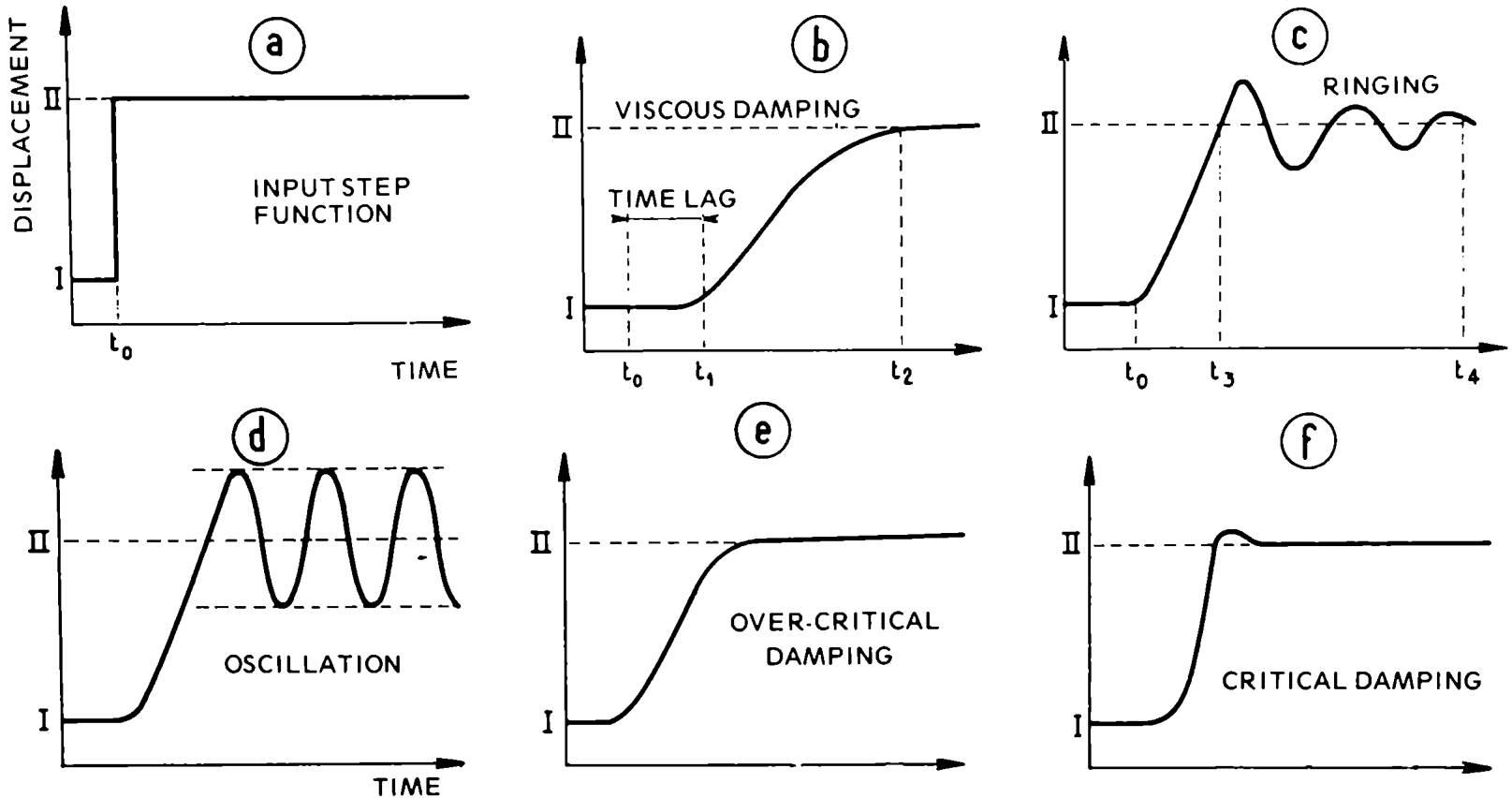


Fig. 1-11. Different types of response displayed by a servomechanism to an input step function.

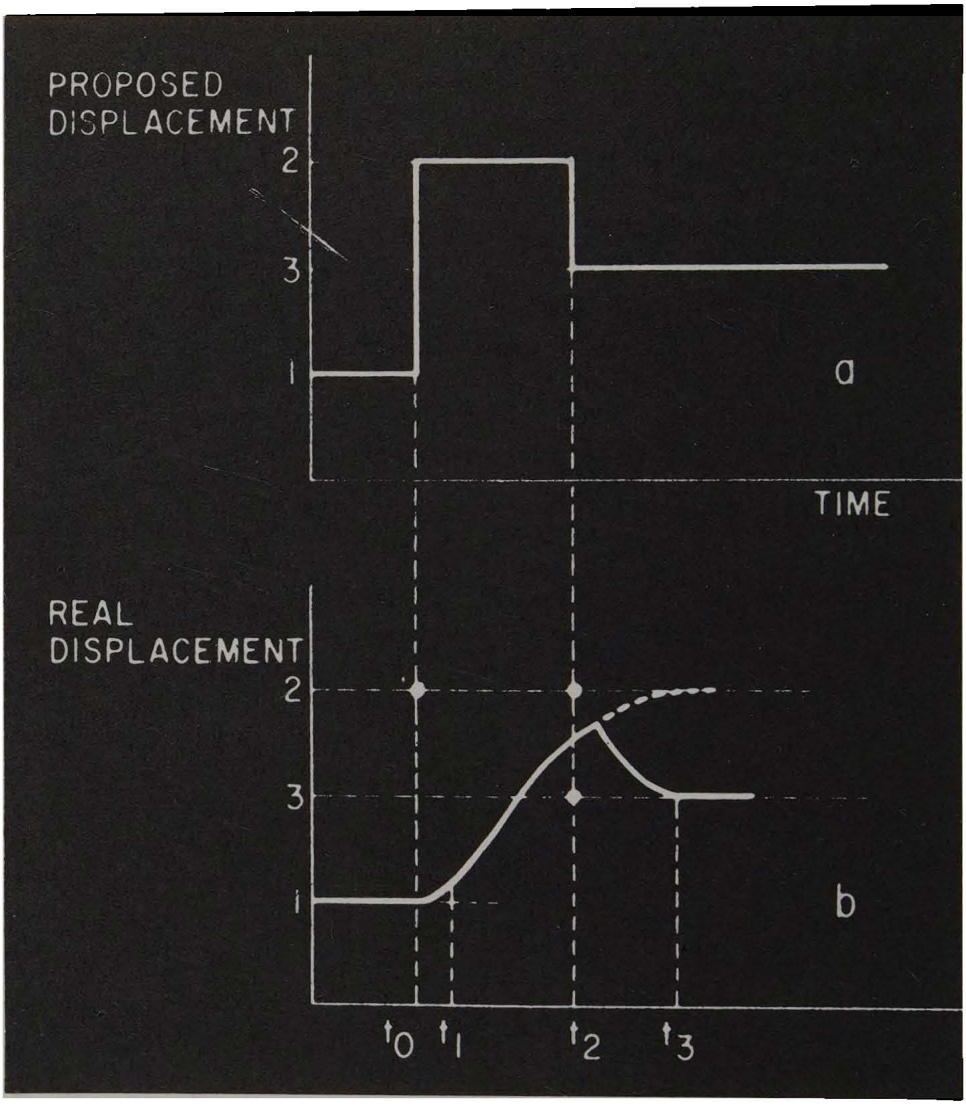


Fig. 1-12. A slow system (B) can only approximate the proposed displacement (A).

some systems are designed for a displacement speed proportional to the error-signal amplitude. This results in a characteristic such as shown in Fig. 1-13.

To understand why a servomechanism is unstable and to find corrective measures, first consider the amplifier in Fig. 1-14 with part of its output fed back to its input. The behavior of this device depends on the input-output phase relationship. If the input and output are out of phase, part of the output voltage fed back opposes the input voltage. This is **inverse feedback**, which reduces overall distortion, improves gain stability, reduces phase shift and widens the frequency response band. On the other hand, overall amplifier gain is reduced. Such an amplifier is inherently stable between its upper and lower

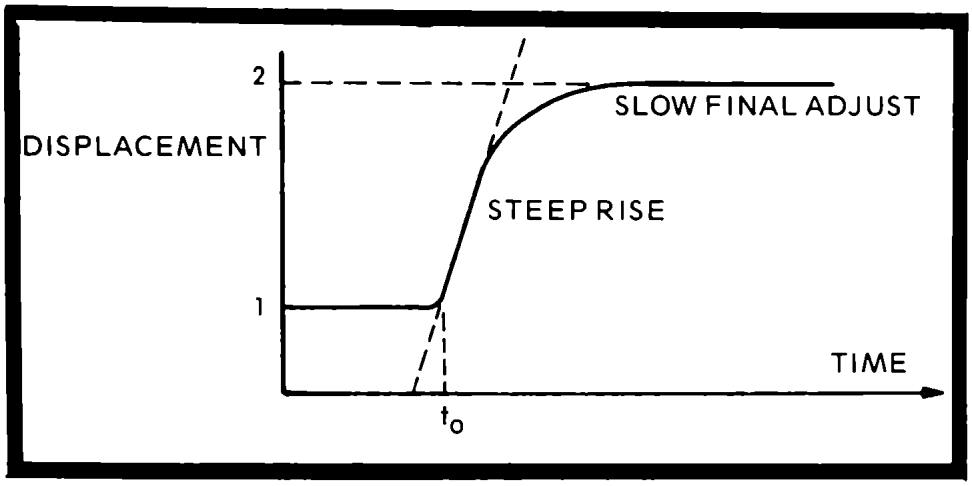


Fig. 1-13. Curve showing displacement speed in proportion to signal amplitude.

frequency limits. In the vicinity of the roll-off regions, the gain decrease is accompanied by a phase variation of the transmitted signal. If this variation attains 180 degrees, the signal fed back is in phase with the input, reinforcing the output quantity; the amplifier builds up oscillations or, at best, becomes unstable, with a somewhat enhanced gain. This can be put quite generally:

"A system is stable if its global open-loop phase difference is less than 180 degrees for the roll-off frequencies."

If this condition is not fulfilled, there will be damped or even suspended oscillations. In practice, a close approach to this limiting condition is carefully avoided, and some security margin called **phase margin** of about 45 degrees is retained.

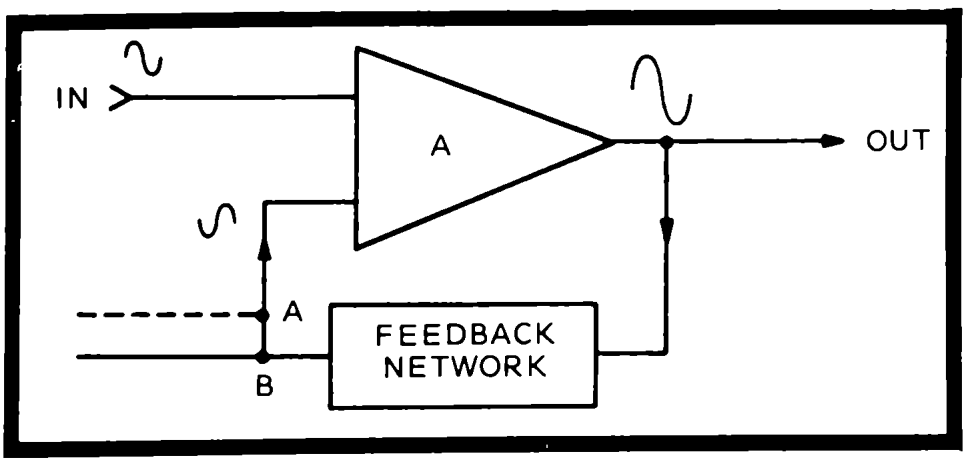


Fig. 1-14. This diagram shows the phase relationships in a feedback amplifier.

To trace the gain and phase characteristics of such a system, gain G (in db) is plotted vs frequency f on logarithmic scales (Fig. 1-15). Just below is traced the phase ϕ variation for the same frequencies. The G characteristic crosses the frequency axis at M (no gain equals 0 db); this is the cutoff frequency f_0. On the ϕ characteristic, f_0 corresponds to a phase angle between 180 and 240 degrees and a **phase margin** of m. Being situated below the 180-degree axis, this segment is counted negative (m is less than 0). On the other hand, point N crossing the 180-degree line corresponds to a **gain margin** n greater than 0. These two disparities indicate instability. Here is a major criterion to be retained:

"To be stable, a system has to have a positive phase margin and a negative gain margain."

An unstable system can, however, be made stable by means of correcting devices such as the RC network in Fig. 1-16, which looks like half a Wien bridge. The gain and phase charac-

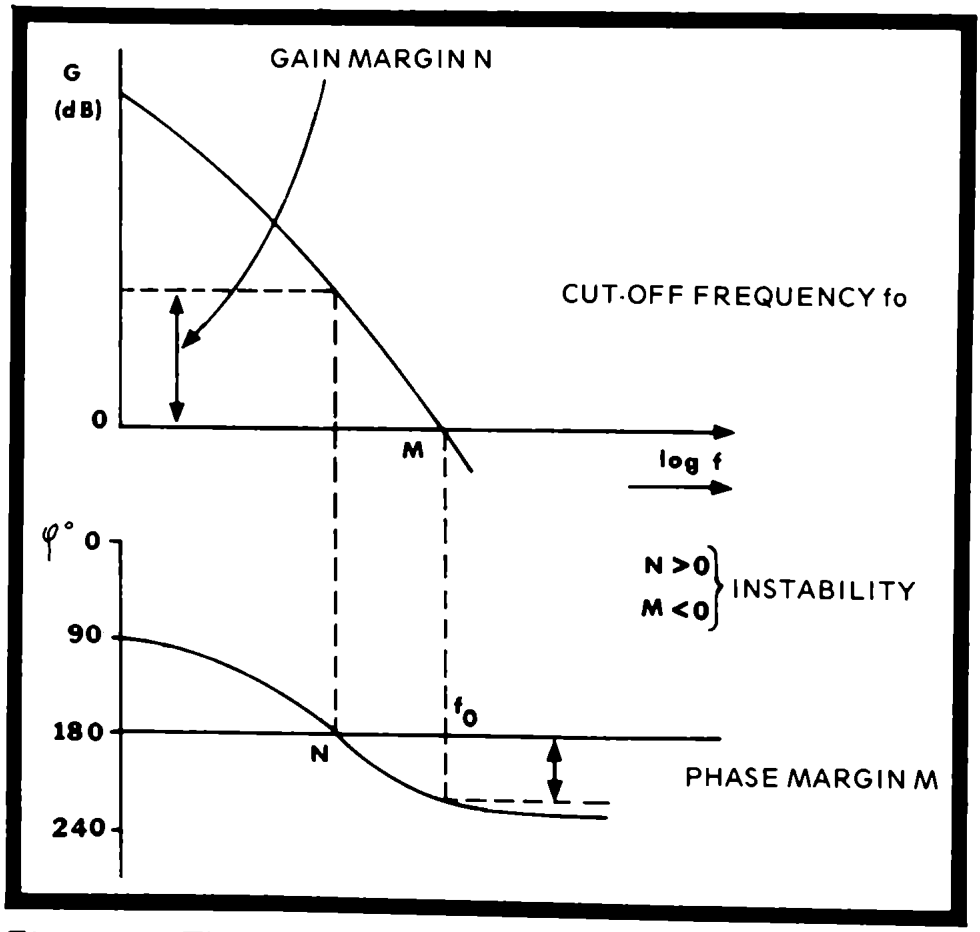


Fig. 1-15. The result of a positive gain margin and a negative phase margin is instability.

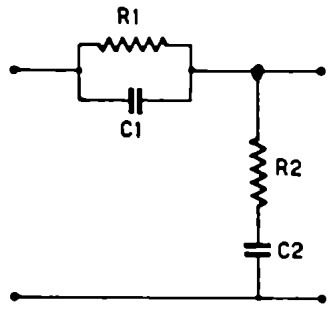


Fig. 1-16. A network such as this is used to correct instability.

teristics G' and ϕ ' of this network are shown in Fig. 1-17. The logarithmic gain scales allow for adding G and G', and phases ϕ and ϕ ', of course, can be added, too. This has been done in Fig. 1-18, where the heavy line represents the corrected gain and phase characteristics, featuring a negative gain margain n' and a positive phase margin m'. Due to the correction device, the system now is stable.

Instrument makers offer sophisticated automatic tracers of different types which produce stability-condition indicating diagrams. However, it is not too difficult to measure G and ϕ for different frequencies using average laboratory equipment. A very low frequency (VLF) generator covering from, say, 0.01 to 60 Hz is needed to feed the system's input. The feedback loop is broken, and Test Points A and B thus obtained are connected to the X and Y inputs of an oscilloscope. The vertical amplitude of the pattern then represents the gain, and the input-output phase relationship can be obtained from the phase ellipse for the applied frequency. There are special

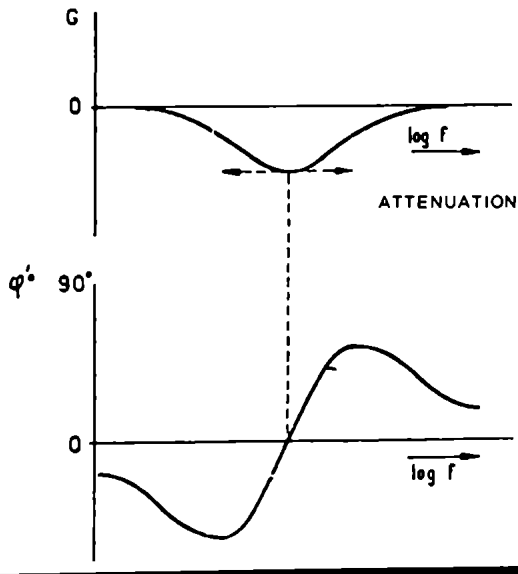


Fig. 1-17. Curves illustrating the attenuation and phase characteristics of the network in Fig. 1-16.

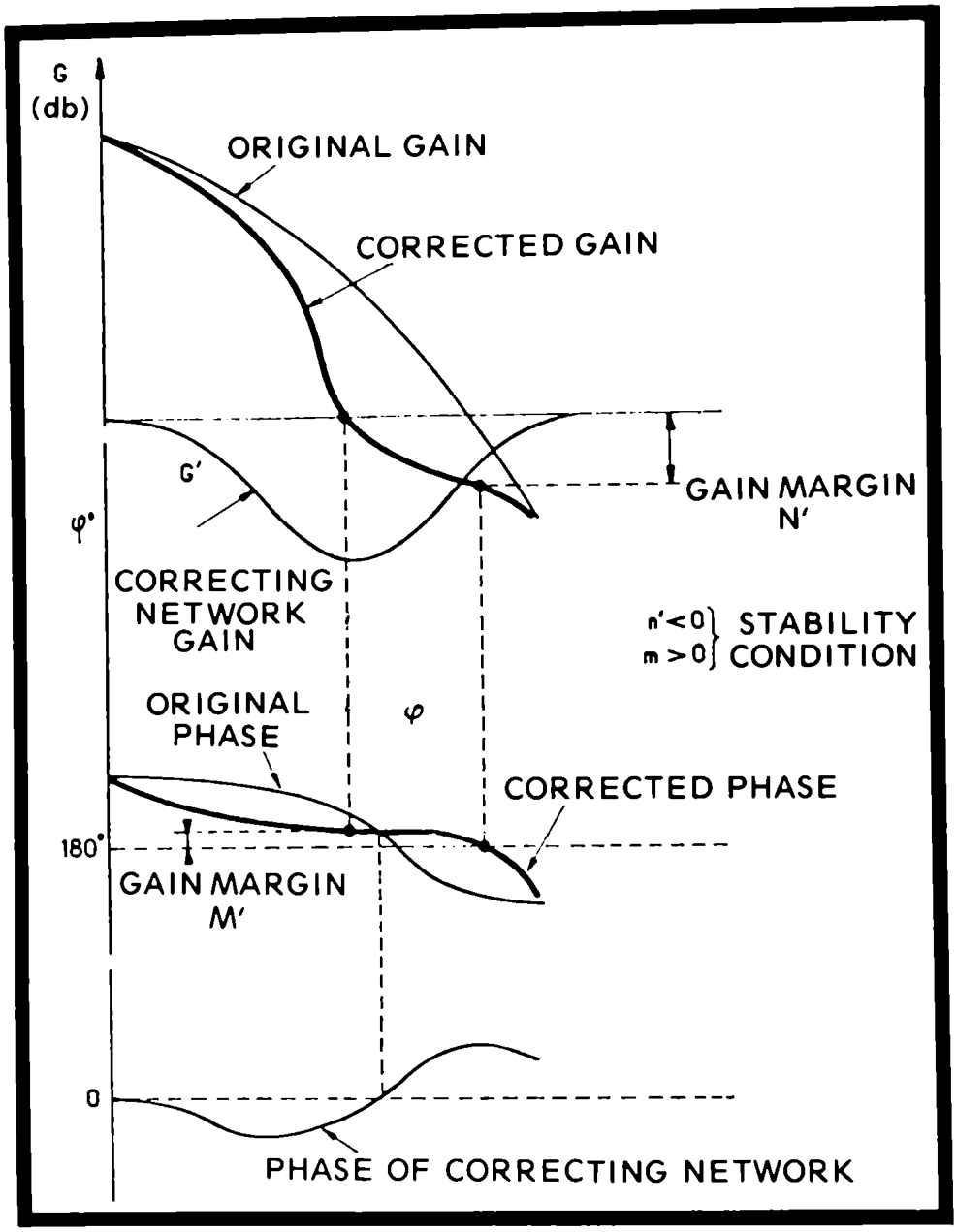


Fig. 1-18. Gain and phase characteristics of a network stabilized system.

"transferometers" designed for this application; they contain calibrated attenuators and phase controls, making for easy reading. However, in some cases it is not practicable to open the control loop of an amplifier in order to study the gain and phase characteristics. Unfortunately, such a straight forward processing is not possible when dealing with the powerful drives used in steel mills, for instance.

A servomechanism can be stabilized, too, by mechanical rather than electric means. (Anyway, the control loop contains mechanical elements with characteristics such as

inertia and friction that are much like electric variables such as inductance and resistance.) Viscous damping of a rotating element can be obtained by a small magnet which induces eddy currents in an electrically conductive disc connected solidly with the rotating axle; this device is used in electricity meters to stop the sensitive movement upon an interruption in current flow (Fig. 1-19). Another damping device (Fig. 1-20) is composed of a flywheel freely rotating on a shaft and a disc connected solidly with the shaft. The flywheel is driven by an oil-impregnated felt pad located between the disc and flywheel. Because of the inertia of the flywheel, this device opposes quick speed variations, but its action is somewhat lessened by the viscous drive. The dash-pot in Fig. 1-21 uses a leaky plunger gliding in an oil- or air-filled cylinder to damp linear displacements.

NULL METHODS FOR ACCURATE MEASUREMENT OF PROCESS VARIABLES

Industrial controls always imply accurate measurement of a magnitude critical for optimum processing. A rather high accuracy (percentwise) is frequently required, and it must be

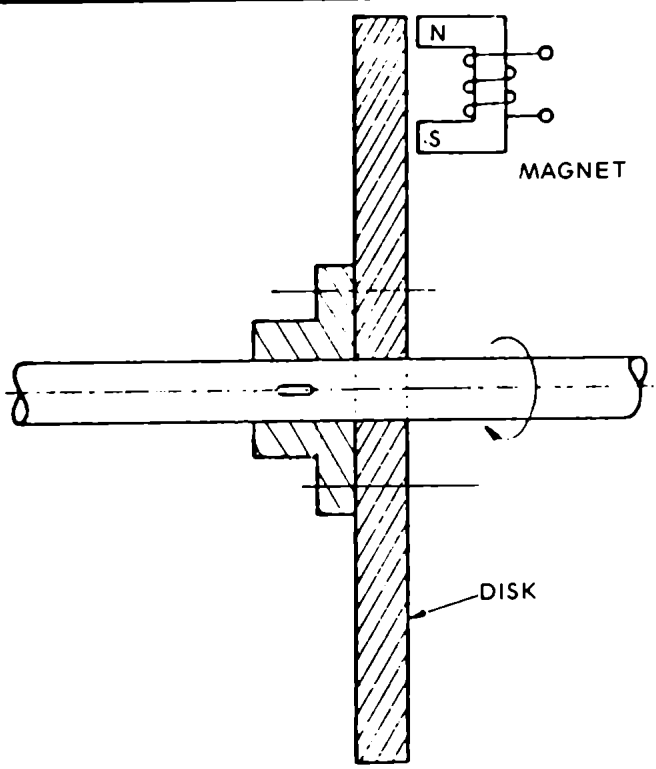


Fig. 1-19. Diagram of an eddy current damping device.

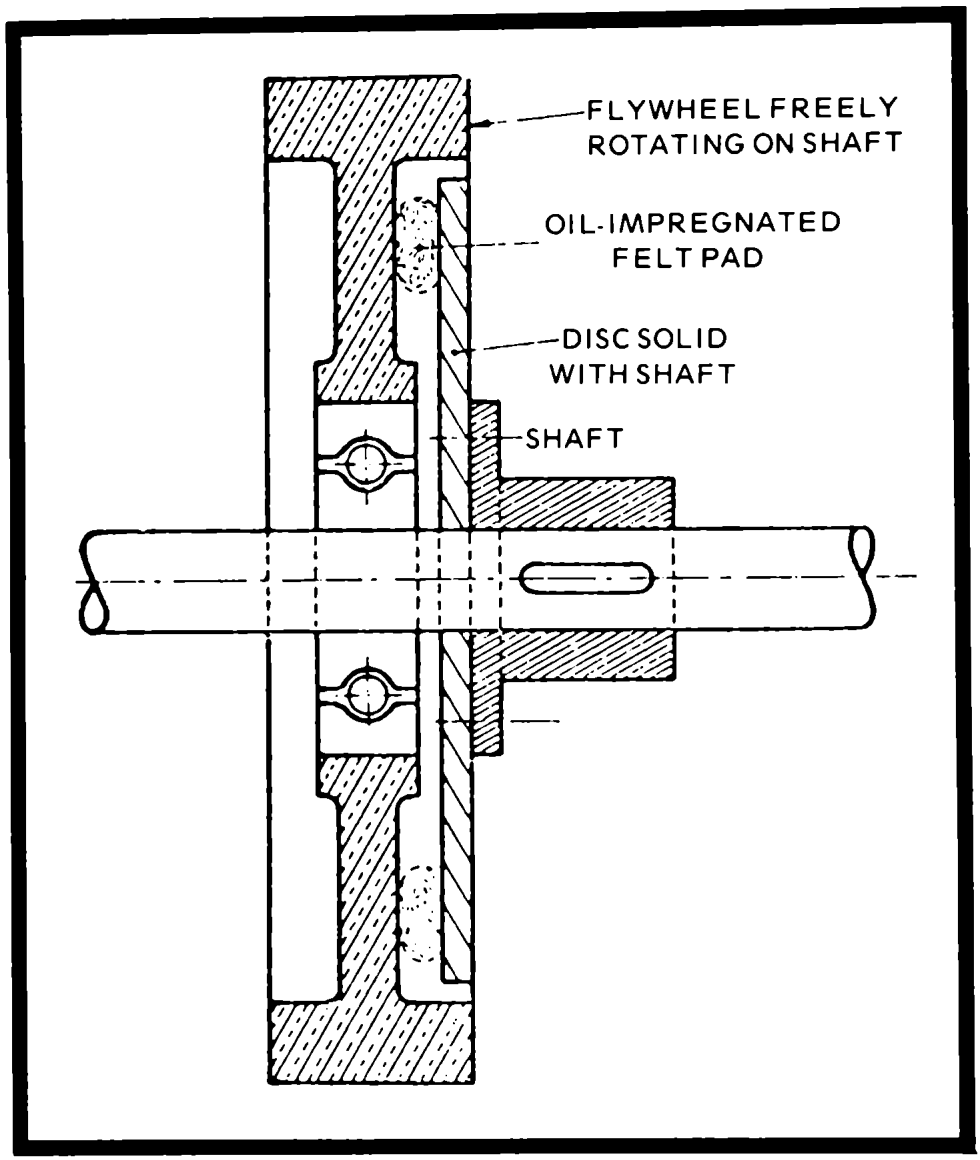


Fig. 1-20. Drawing of a viscous friction damping device.

obtained with a sturdy transducer often working in a rough environment. Thus, the use of delicate and fragile laboratory instruments cannot even be considered. Therefore, methods must be devised and transducers designed to grant accurate and consistent measurements in spite of most adverse conditions.

An unknown quantity can be evaluated or measured by comparison with a known quantity of the same magnitude or a standard that is varied until both are equal. A null detector is required to display this equality. This is the operating principle of a beam-balance that compares one weight to another; the "null" at balance is displayed by the pointer. While handling such a beam-balance is time-consuming if accurate

weighting is required, there are spring-loaded balances for quick evaluation. The magnitudes to be compared are dissimilar now, for the unknown weight is compared to the elongation of a calibrated spring. The familiar D'Arsonval meter is just another example of such a comparison of dissimilar ("analog") quantities, for the rotation of the moving coil caused by the current to be measured is opposed by the torque of the fine spirals used as a reference, and the pointer indicates the balance point on a volt- or ampere-calibrated scale (Fig. 1-22).

As noted above, the relative (percentwise) variation of the quantity measured frequently is very small in process control systems. An analog method such as the d'Arsonval meter features an accuracy of about 1 to 5 percent, and some 0.2 percent at best, using expensive laboratory instruments unfit for industrial use. On the other hand, much higher accuracies are reliably obtained with sturdy industrial transducers using null methods such as the beam-balance. In an electric null network, the balance condition is detected by the **absence** of any current, and the sensitivity of the system to very small changes can be increased using a null detector incorporating a high-gain amplifier. Furthermore, this principle allows for designing automatic null-seeking and follower devices. Two basic circuit types are used: **bridges** for balancing impedances, and **potentiometers** for opposing voltages or currents.

Bridges

All bridge circuits are derived from the Wheatstone bridge, composed of four ring-connected resistors X, R, a and

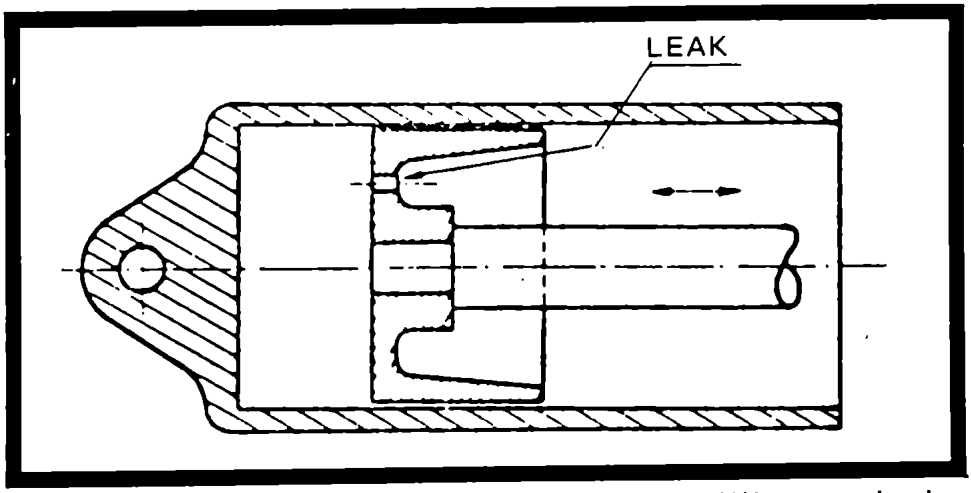


Fig. 1-21. A dash-pot damping system utilizes a leaky piston.

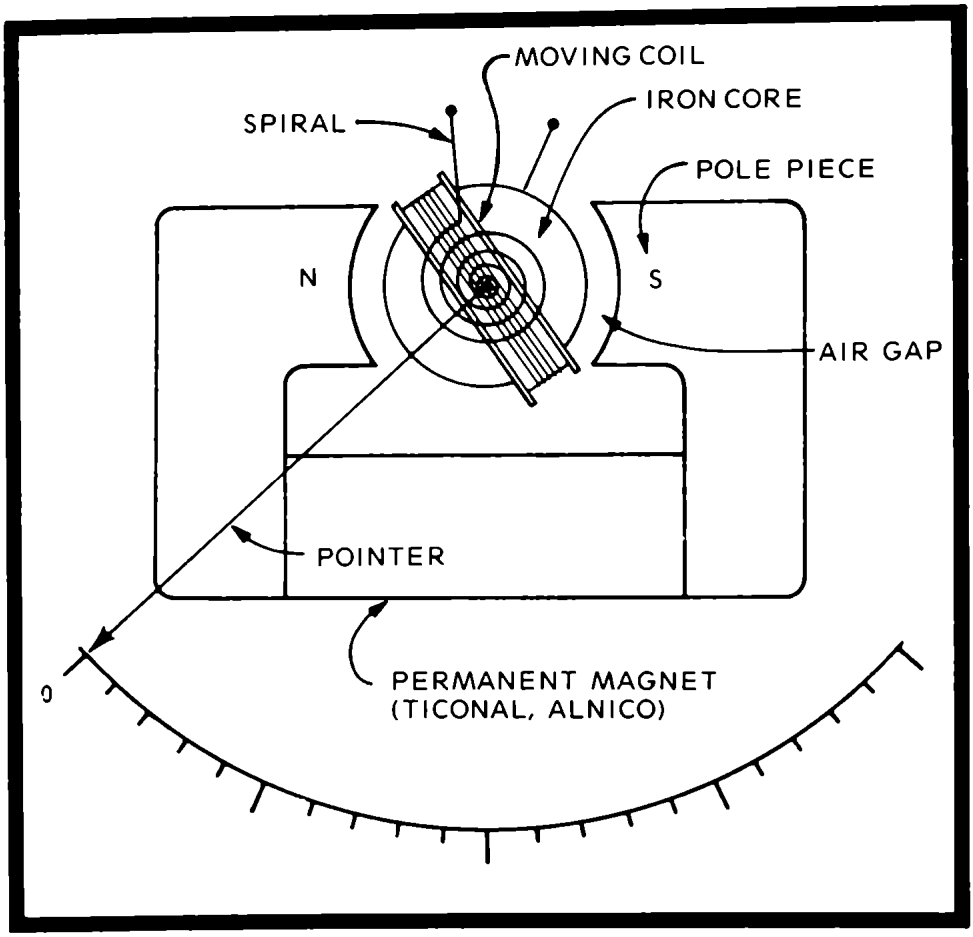


Fig. 1-22. Drawing of a d'Arsonval moving-coil meter.

b in Fig. 1-23. One diagonal of the quadrilateral thus obtained contains a battery, and the other a null detector (a galvanometer in DC-fed bridges). It can be easily shown that the bridge is balanced (no pointer deflection, or null) if:

$$\frac{a}{b} = \frac{X}{R}$$

or

$$X = R\,\frac{a}{b}$$

Thus, unknown X equals standard R times the ratio a over b. There is no need for knowing the actual values of a and b, only their ratio. In the commercial bridge in Fig. 1-24, fixed ratios of 1, 10, 100 and 1,000 are available, as well as one-tenth, one-one hundredth and one-one thousandth by permuting a and b. In another type of bridge, which is quick-reading rather than

ultimately accurate, ratio a over b is formed by a potentiometer calibrated to read a over b directly from, say, 1 to 10. Both types of variable ratio arms are conceived to minimize the number of (costly) standard resistors required to cover a given resistance range. In industrial control, extensive variations generally are not to be expected.

Notice that the error voltage decreases the more a over b differs from unity ratio, so that the highest accuracy for a given null detector sensitivity is obtained with a equal to b. If X stands for a thermocouple used as a furnace temperature transducer or a strain gauge for elongation measurement, accuracies from 0.1 to 0.001 percent may be required to match reading accuracy requirements.

The same bridge can be supplied by an AC source such as an oscillator or a main power line transformer winding; the galvanometer is replaced by a suitable AC detector, generally preceded by an amplifier to increase its sensitivity. Provided stray capacitances are negligible (if they are not, compensation is possible), the same balance condition holds.

If one or more elements of the AC-fed bridge are not "pure" resistances, but capacitances, inductances or devices presenting resistive as well as inductive or capacitive components also, we have an **impedance bridge**. There are many

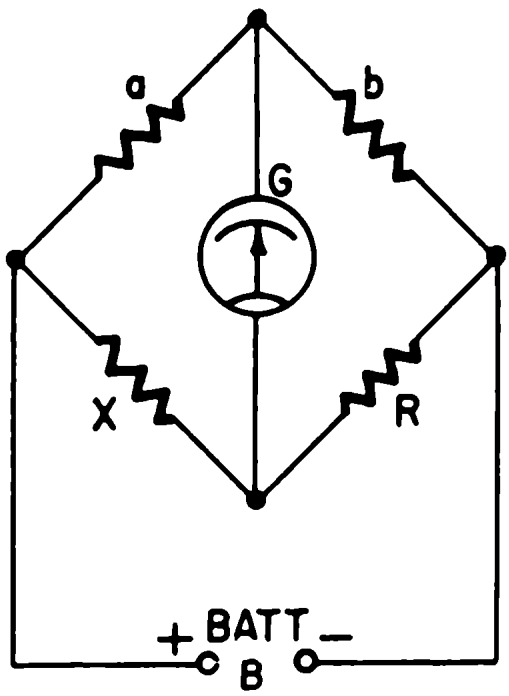


Fig. 1-23. Circuit of a basic Wheatstone bridge.

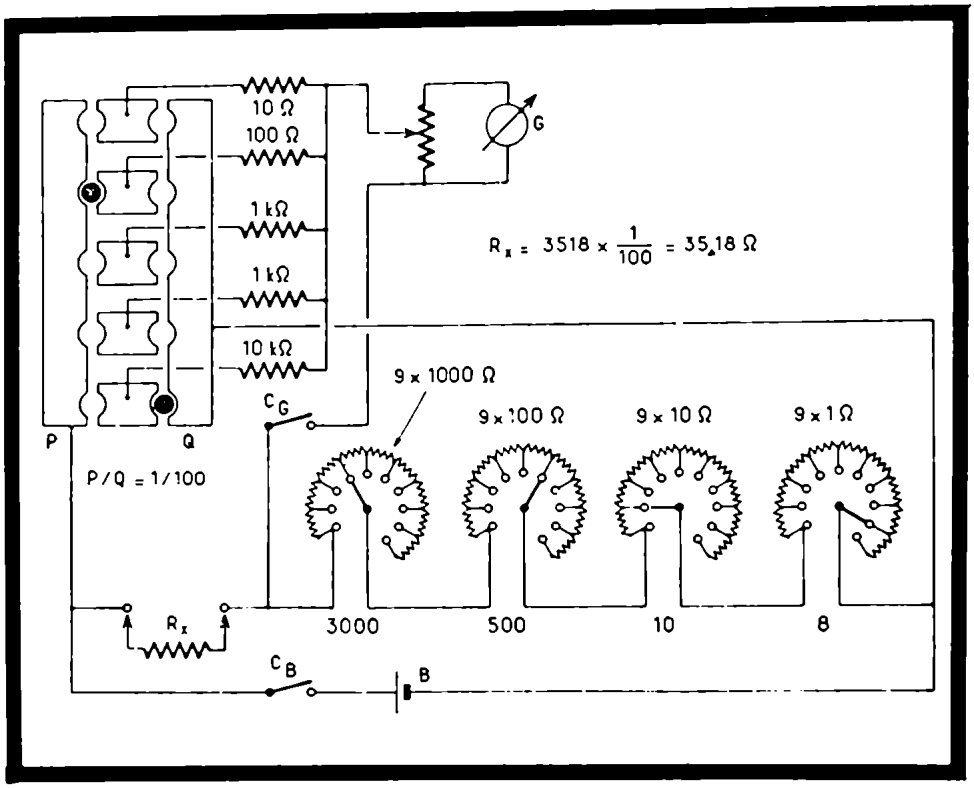


Fig. 1-24. Schematic of a precision Wheatstone bridge.

such bridges, depending on the actual disposition and the nature of the bridge network components, and an unknown reactance can be balanced by a like as well as by a "complementary" reactance—a capacitance by another capacitance or an inductance by a capacitance. The problem is dealt with in a number of textbooks. For industrial control-minded readers, it is important to remember that in an impedance bridge there are two balance conditions to be satisfied: it must be balanced for the resistive as well as for the reactive (capacitive or inductive) components. Hence, there are two independent (though interacting) balance controls to be handled to obtain a null. In industrial control, however, the transducer deals with comparatively small variations, so that only one component can be balanced; the unbalance of the other is negligible.

POTENTIOMETER METHOD FOR VOLTAGE MEASUREMENT

The moving coil meter in Fig. 1-22 allows for quick voltage or current checking, but its ability to display small changes is inadequate for industrial applications. Here, too, a null method is preferred, for it features a potentially much higher

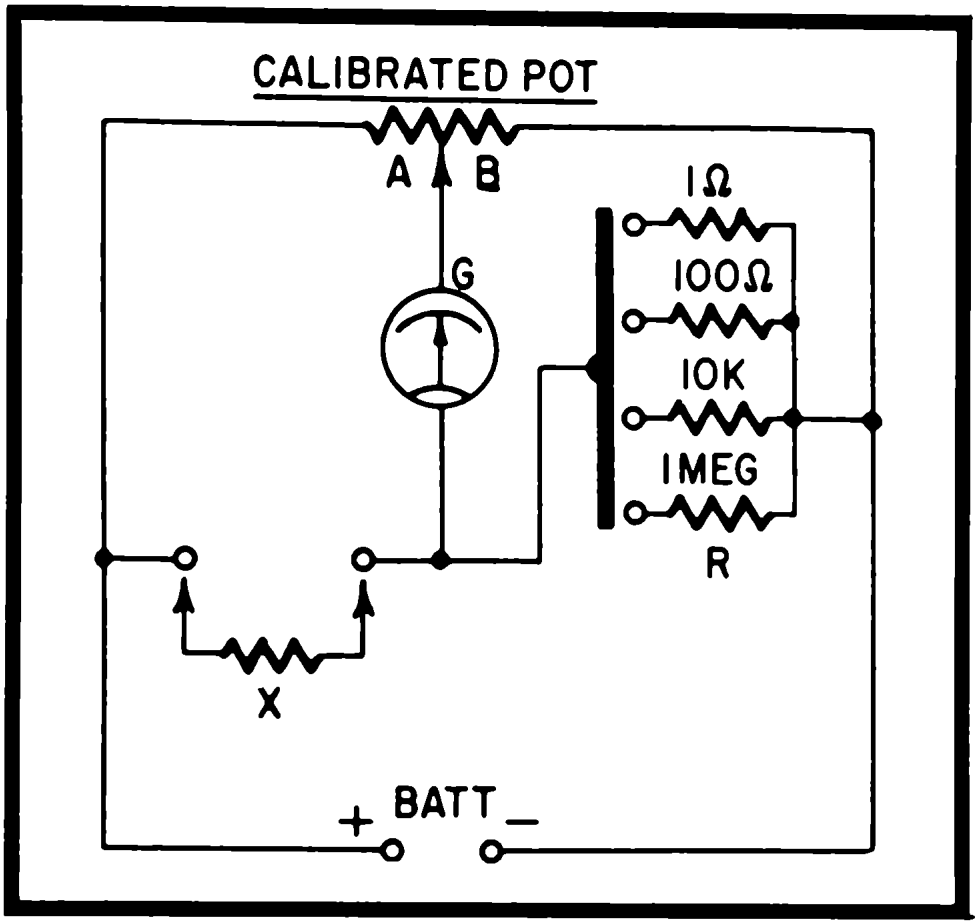


Fig. 1-25. Quick-reading potentiometer-type bridge circuit.

accuracy and, still more important, it readily leads to automatic null-seeking controls.

The principle of potentiometric voltage measurement is illustrated by Fig. 1-26. An unknown voltage Eu is opposed to a

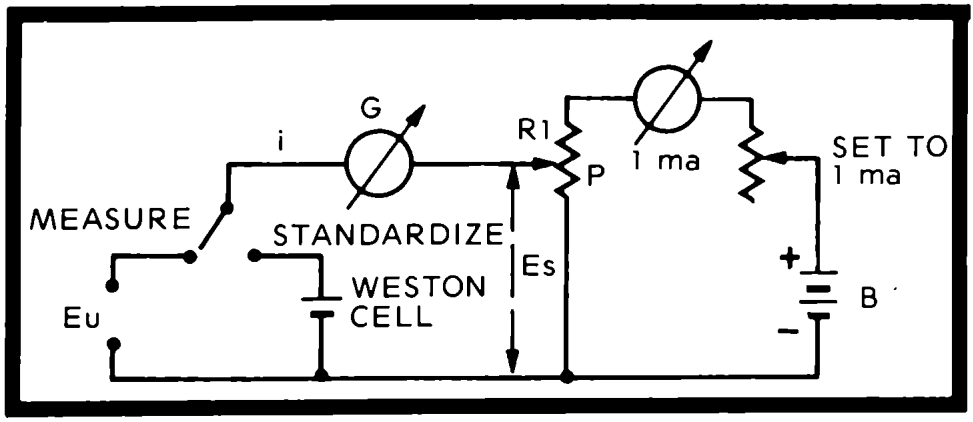


Fig. 1-26. This circuit illustrates the potentiometric method of voltage measurement and comparison.

standard voltage Es made up of the voltage drop across the R1 portion of the calibrated potentiometer P (between the slider and ground). By a battery B, P is supplied with a metered variable current that can be set to a calibration point of, say, 1 ma, developing a potential difference of 1 mv across every ohm of resistance of R1. Polarities are chosen to oppose Eu and Es, so that current i in galvanometer G becomes zero when Eu equals Es. As previously, the reading accuracy can be greatly increased by amplifying error current i. This device can be made automatically null-seeking and null-following, too.

A voltage source such as battery B does not feature a stability comparable to that of a resistance network; means for standardizing may, therefore, be mandatory. A Weston cell yielding an open-circuit voltage of 1.0185v, plus or minus 0.01 percent at 68 degrees F can be switched across the input. As the calibrating voltage drop across R1 is defined by 1 mv per ohm, R1 is set at 1,018.5 ohms, and auxiliary current Is is adjusted for i equals 0. This calibration holds as long as Is remains unchanged. The Weston cell, delicate to handle, is nowadays often replaced by a temperature-stabilized precision zener diode network featuring about the same stability and immunity to shorts, temperature changes, shock and vibration. Precision regulated low-power supplies are readily available, too, to replace the battery.

In industrial applications, the standardizing voltage source is sometimes supplied by a reference transducer similar to the measurement transducer, but subjected to the physical reference conditions. This may be, for instance, a thermocouple held at a reference temperature.

The potentiometric method applies to AC voltages, too, provided they are of the same frequency and phase. Voltages of difference frequency and varying phase, however, can be compared using the rectified DC voltages. As rectifiers have a parabolic segment at the bottom of their characteristic, no precise rectified voltage comparison can be expected at low amplitudes (less than several volts). The high accuracy of four to five decimal digits featured by modern digital voltmeters has only been made possible by the potentiometric method (see Chapter 4). Yet there is no potentiometer, be it manual or motor driven, with the capabilities of a slope generator producing a uniformly rising voltage like a time base, in which the coincidence of both voltages is detected by a comparator circuit. A high stability reference source of the zener type is self-contained.

CHAPTER 2

Transducers

The term "transducer," in a broad sense, means "a device for converting one form of energy to another." In control engineering, the "other" form of energy is an electrical quantity, the only message an electronic circuit can understand. This signal informs the control system of the actual situation characterized by some physical, chemical or other parameter to enable it to act upon a process or an operation. The transducer output signal has to be proportional to the magnitude of the input, whatever form it may take—displacement, pressure, temperature, light, and other.

For every application, a suitable sensing device exists or can be designed, and there even is a choice to be made between different solutions; the number of possible transducers, therefore, is enormous. The operating principles, however, are limited to a few basic patterns. The devices are classified here according to the electrical quantity in which the variations are related to the magnitude to be sensed. For a given input variation, a transducer is designed to yield the greatest possible output variation with a low noise level for reliable operation of the control system. Obviously, its operation has to be stable, reproducible and reliable.

VARIABLE RESISTANCE POSITION TRANSDUCER

Resistive transducers are connected into DC- or AC-fed Wheatstone bridges. As a potentiometer can be considered forming two ratio arms, the remaining bridge network can be made up of only two resistors, or even a second potentiometer (Fig. 2-1). A potentiometer can be used as an angular position transducer where the output magnitude is resistance R1 or R2, or even the ratio R1 over R2. Rotating an axle attached solid with the potentiometer shaft will unbalance the bridge formed by the two potentiometers and deflect the galvanometer needle. The null can be restored by varying ratio R3 over R4; this second potentiometer, thus, can conveniently be calibrated to indicate angular displacement of the first one.

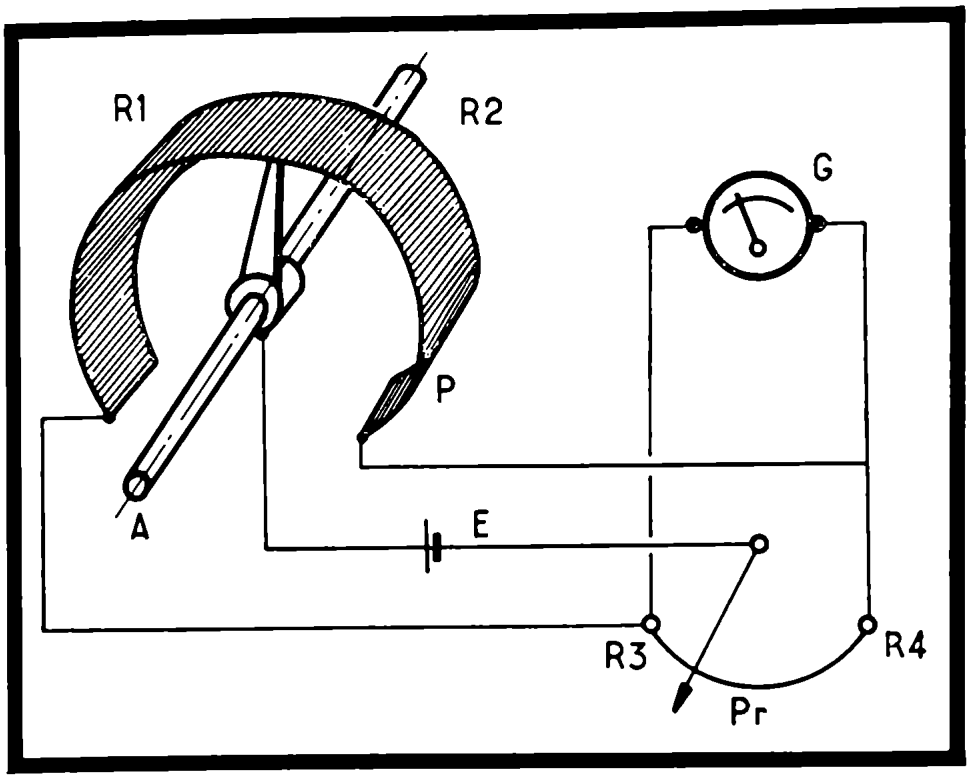


Fig. 2-1. Angular position transmitter output is applied to a bridge-connected potentiometer.

The conventional one-turn potentiometer offers a rotation angle of about 300 degrees, but there are 10-turn potentiometers covering 360 degrees. A much larger or smaller angular displacement can be geared down or amplified mechanically using a suitable gearbox, Multiturn pots allow for a very high angular definition (0.01-percent) and, therefore, are preferred for accurate position sensing. These devices are composed of a resistive wire uniformly wound on an insulating string of circular cross-section, which is introduced in a helical groove within the cylindrical housing of dielectric material. By means of a screw, the slider is made to follow the helix to contact the wire on the string, hence the term helical pot.

To sense a linear displacement, a rotating pot driven by rack and pinion can be used. The transformation of a linear into an angular displacement, however, can be avoided, using a linearpotentiometer (Fig. 2-2). By displacing the fixed point of a lever linkage, the amplitudes of the slider and monitored device movements can be matched. Because of the unavoidable stops, only finite displacements can be sensed by potentiometers.

Tiny precision pots are also used in other types of transducers where the physical magnitude to be sensed is first

transformed in a related angular displacement, which the pot transforms in a corresponding variation of resistance or resistance ratio. The manometric or barometric device in Fig. 2-3 comprises a pressure-sensitive flat capsule or box with thin undulated walls. In a barometer, this capsule is sealed, and the varying atmospheric pressure results in compressing or depressing the elastic walls. In a manometer, the pressure to be sensed communicates with the inside of the bow, and its outer ambient is maintained invariable to form a reference. In both cases, one wall is held fixed, and the displacement of the other, which translates the pressure variation, is coupled to the pot shaft by gears and linkages providing for mechanical amplification.

In the manufacturing of light-sensitive papers, air humidity must be maintained within close limits if the papers are not to be spoiled. The hygrometer in Fig. 2-4 is used for

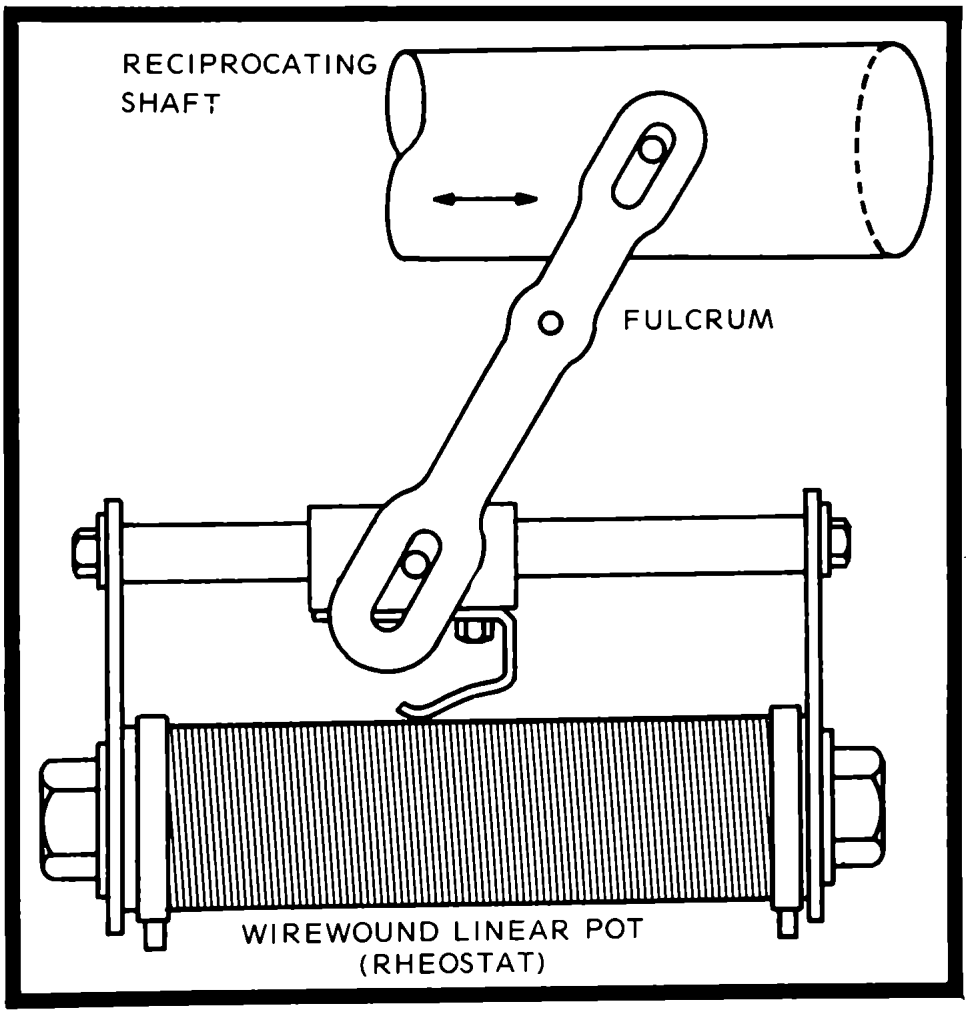


Fig. 2-2. Drawing of a resistive type linear displacement transducer.

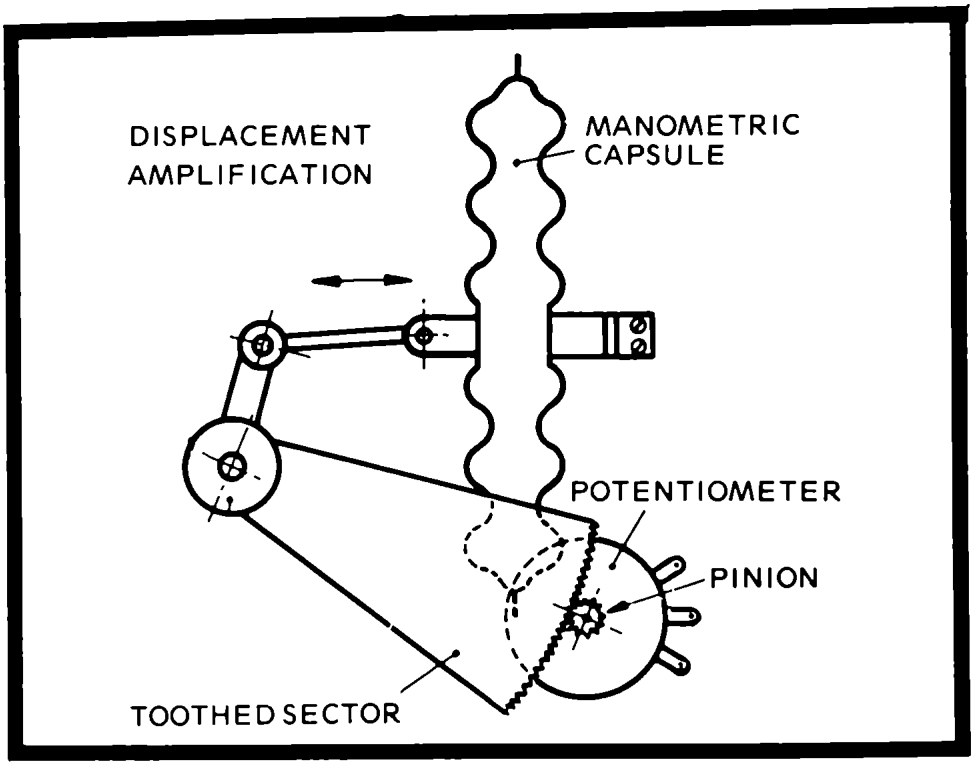


Fig. 2-3. A pressure-sensitive capsule linked to a potentiometer forms a manometric or a barometric transducer.

monitoring humidity. Its "works" comprise a string of human hair that elongates and retracts upon humidity variations as is well known. Up to now, no better or cheaper medium has been found; it seems that blond women's hair works best. Here again, the displacement of the free end of the string is coupled to the shaft of a precision pot by gears and linkages, providing for mechanical amplification.

STRAIN GAUGES

A strain gauge is a transducer which interprets a dimensional variation of a supporting element by a change of resistance. This tiny device (about the size of a postage stamp) consists of a thin piece of paper on which is glued a zigzag length of fine resistance wire (Fig. 2-5). This guage is cemented as perfectly as possible to the structure whose elongation is to be measured. (Without a reliable bond, no accurate results are to be expected.)

Now, suppose a strain is applied to the structure under test, elongating it by a very tiny amount. The gauge paper will be elongated, too, and the bonded wire will be lengthened, somewhat reducing its cross-section. Both of these dimen-

sional changes increase the wire's resistance by a tiny though measurable amount. For reasonably small elongations (well below the elasticity limit of the wire), the proportionality between both variations holds fairly well. Because the wire is well guided in the plastic adhesive that holds it to the piece of paper, compressions can be measured as well as elongations, since the sensitivity is the same.

The relative resistance variation (dR over R) of the gauge of a nominal resistance R in terms of a relative variation of elongation dL over L is given by:

$$\frac{dR}{R} = K\frac{dL}{L}$$

K is the factor characteristic of the gauge sensitivity, generally about 2. For every gauge, the manufacturer indicates K with a 0.5 percent accuracy. Thus, dL over L can be directly calculated from a dR over R measurement. For the proportion to hold, dL over L is generally limited to 0.5 per-

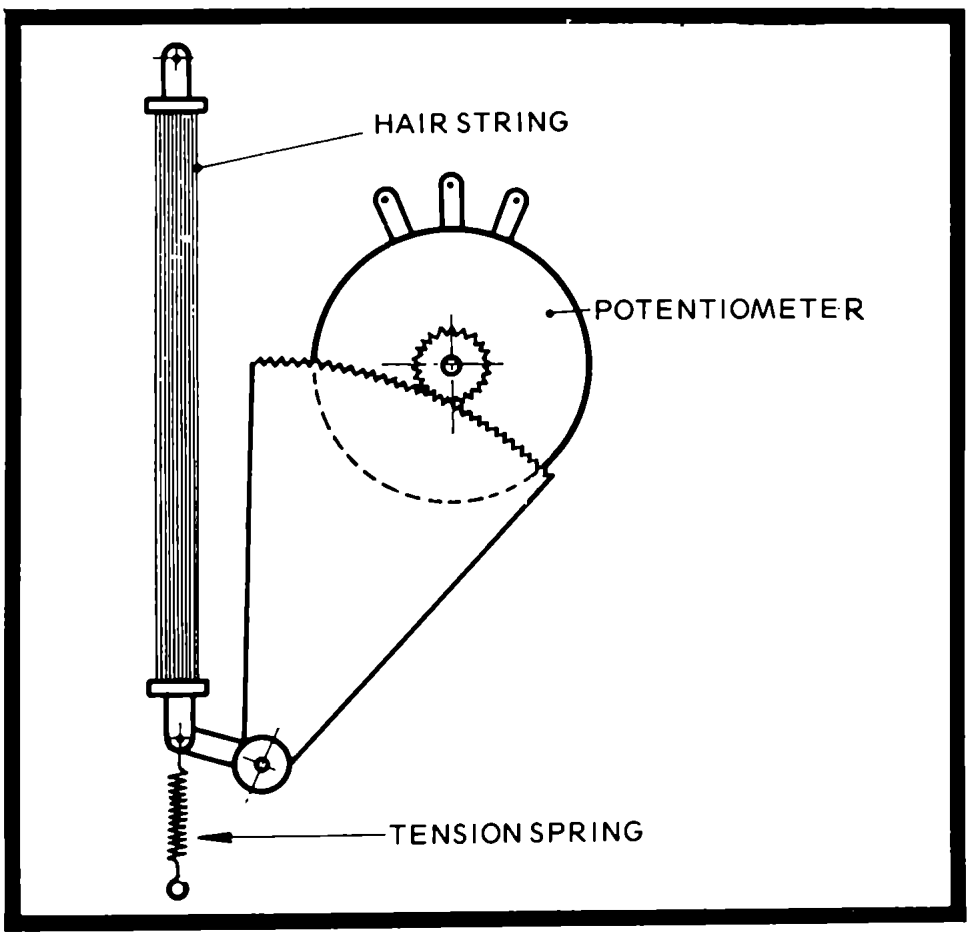


Fig. 2-4. Strands of human hair attached to a potentiometer serve as a humidity monitor.

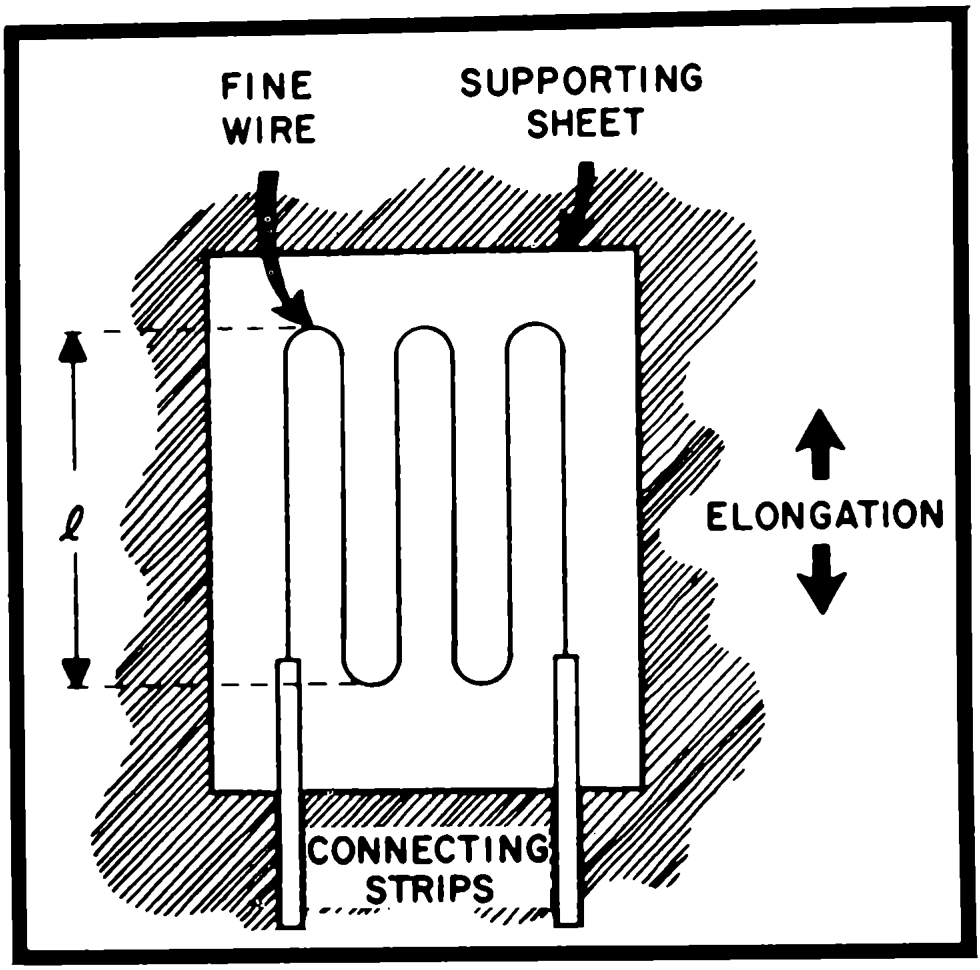
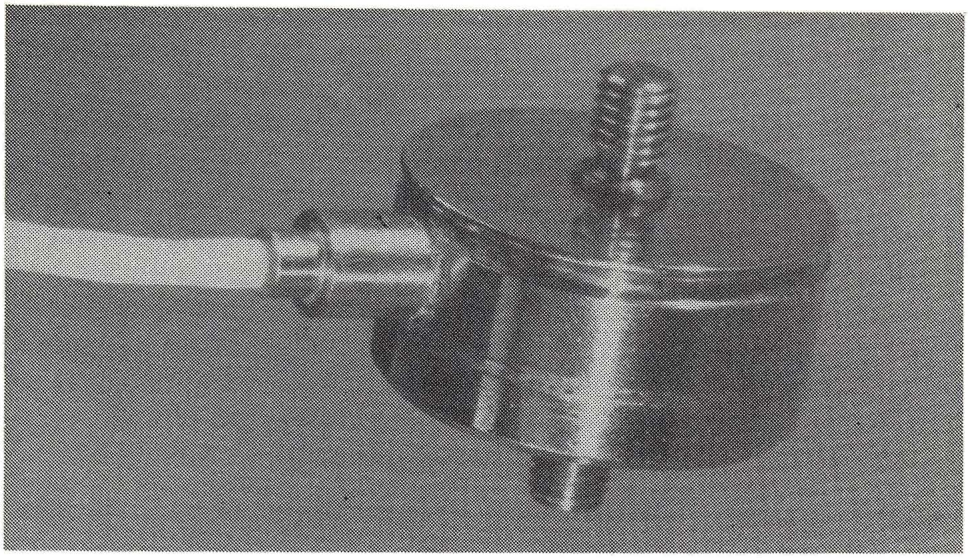


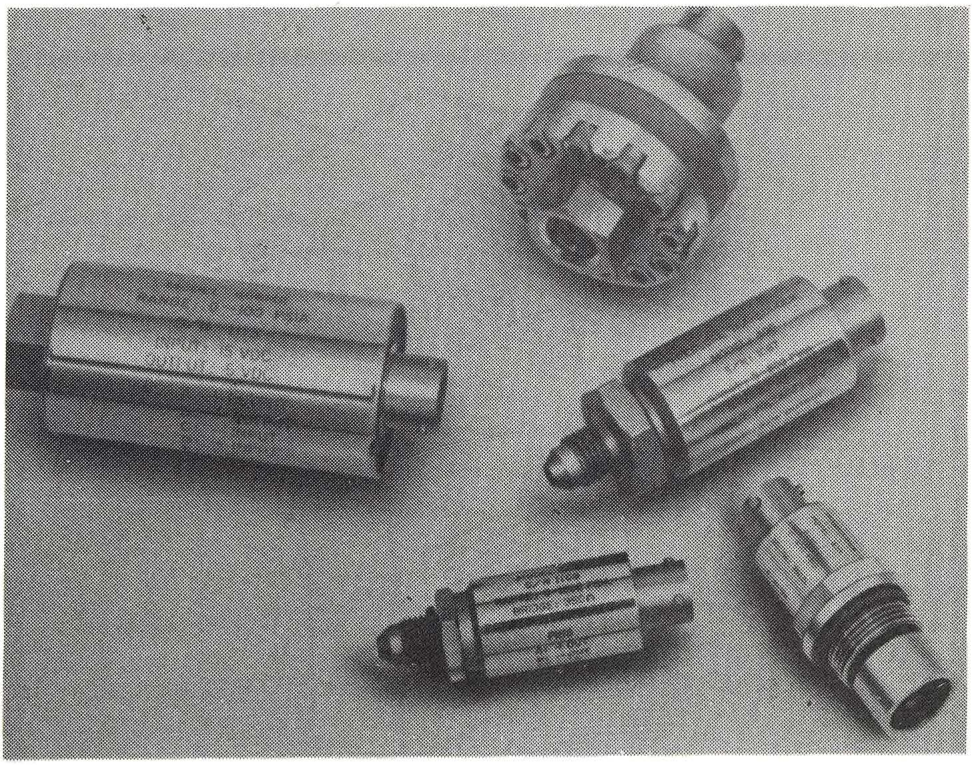
Fig. 2-5. Drawing of a fine-wire strain gauge transducer.

cent. With a K of about 2, this results in an upper dR over R limit of 1 percent, but in actual practice, much lower variations are encountered.

In static measurements, such tiny variations can be measured only with precision Wheatstone bridges, modified to allow for direct dR over R calibration. Moreover, the comparatively small temperature coefficient of a gauge of 10^{-4} to 10^{-5} per degree C can introduce an error of about the same magnitude in a dL over L calculation; therefore, precise compensation is needed. Generally, this is done by connecting a compensating gauge (J2) in the bridge arm adjacent to active gauge J1 as shown in Fig. 2-6. Both gauges are cemented on the same structure to be tested; therefore, both are exposed to the same temperature, thus the temperature effects are canceled. But while the active gauge is placed along the lines of maximum elongation, the compensating gauge is cemented perpendicularly to these lines in a neutral position so that the deformation does not vary its resistance



Tension-compression 30 Series load cell designed to measure forces in a range from 0 to 50 grams up to 0 to 100 lbs. It uses bonded semiconductor strain gauges wired in a wheatstone bridge. Manufactured by Sensatech, Inc.



Five transducers made of welded stainless steel with bonded foil strain gauges. They are designed to operate in temperatures from -65 to +250 degrees F. Manufactured by Consolidated Controls.

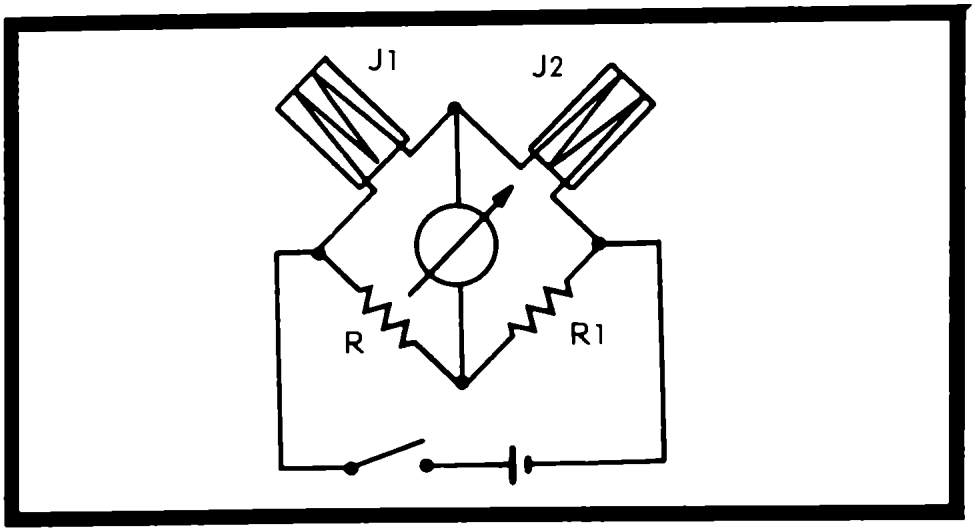


Fig. 2-6. Wheatstone bridge circuit containing active and compensating strain gauges.

(Fig. 2-7). Sometimes the two remaining resistors, R and R', are replaced by strain gauges, too; all four gauges then are active, and the pair of gauges subjected to the same type of deformation are connected in opposite arms of the bridge.

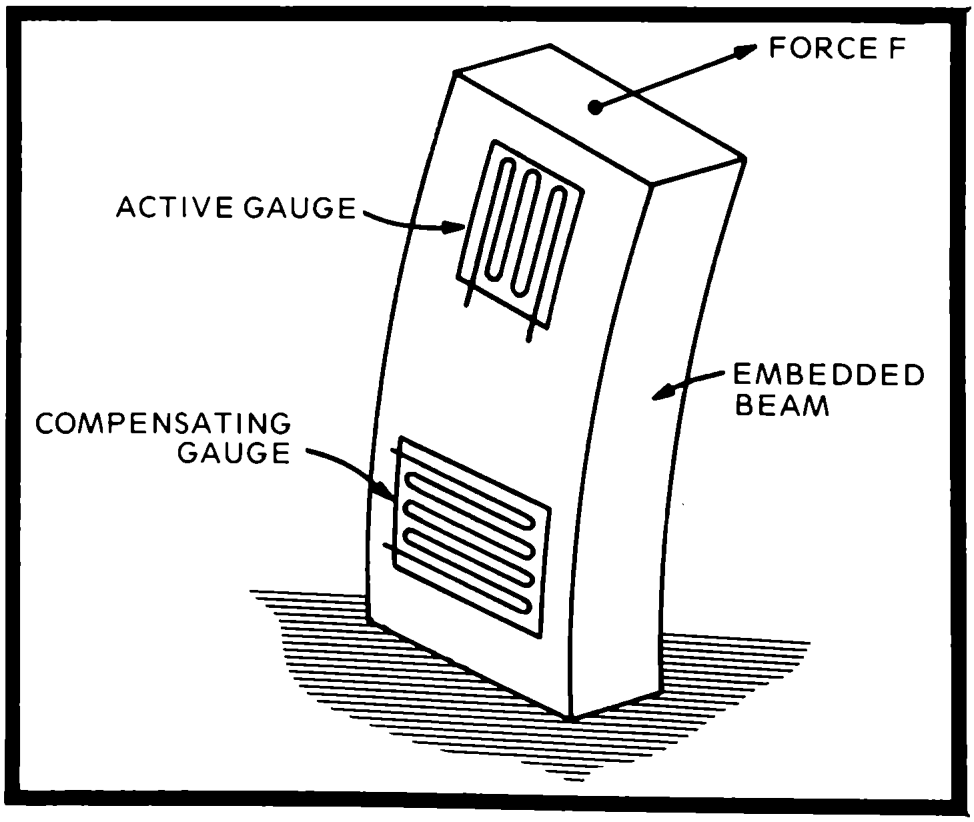


Fig. 2-7. Drawing showing the positioning of active and compensating strain gauges on a beam.

A bridge specially designed for strain gauge measurements is represented in Fig. 2-8. Across the diagonal AC, series-connected resistances are tied to one-pole multidirectional switches K1, K2 and K3 allowing for dR over R variations of 10^{-2}, 10^{-3} and 10^{-4} per step, respectively, and P makes for ultimate adjustment.

Transducers for various magnitudes can be constructed using strain gauge techniques. For instance, to measure the pull exerted by a locomotive hooked to pull a train, two active gauges, J1 and J2, are cemented on the traction bar in the direction of the effort, and two passive (compensating) gauges, J3 and J4, are placed perpendicularily to the pull axis. (Fig. 2-9). The same device can be used to measure compression. There are commercial "load cells" for weighing heavy loads using the same principle. With three such load

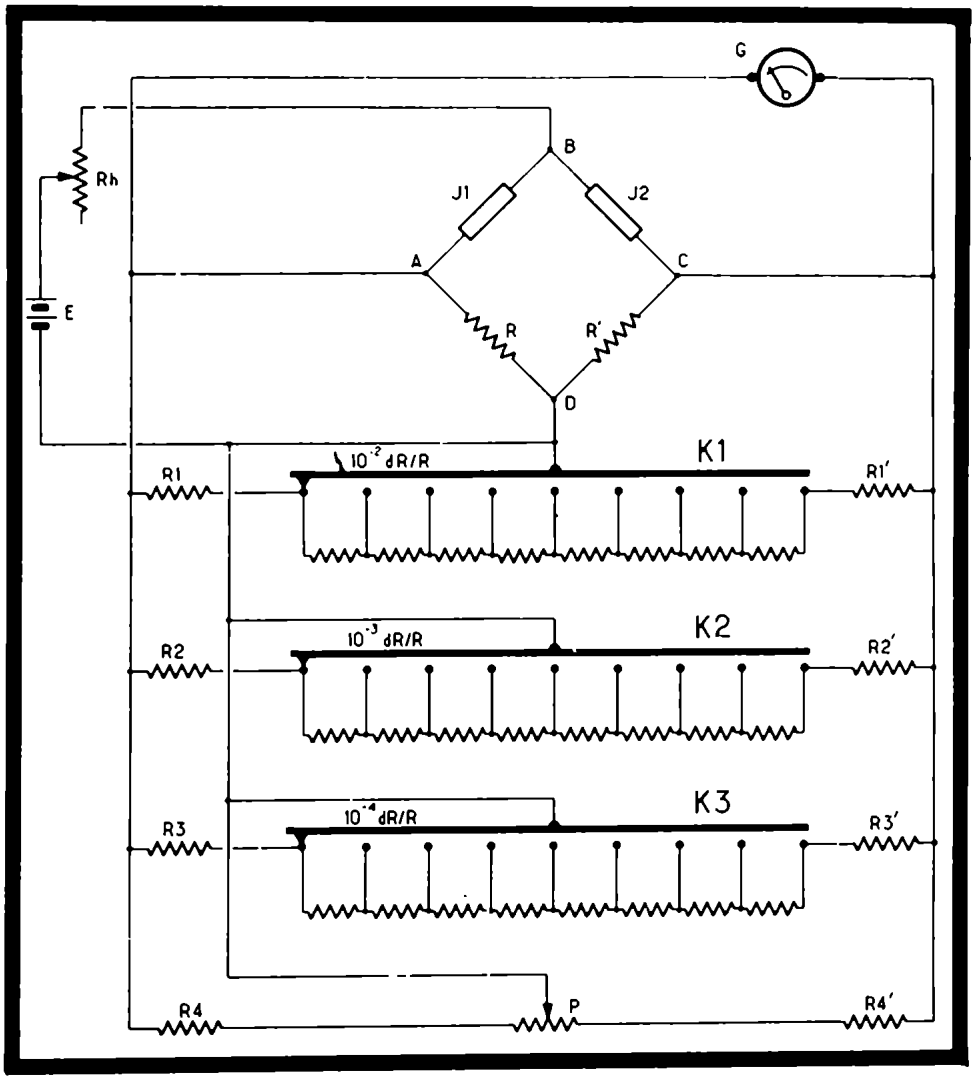


Fig. 2-8. Schematic of an extensometer bridge for static measurements.

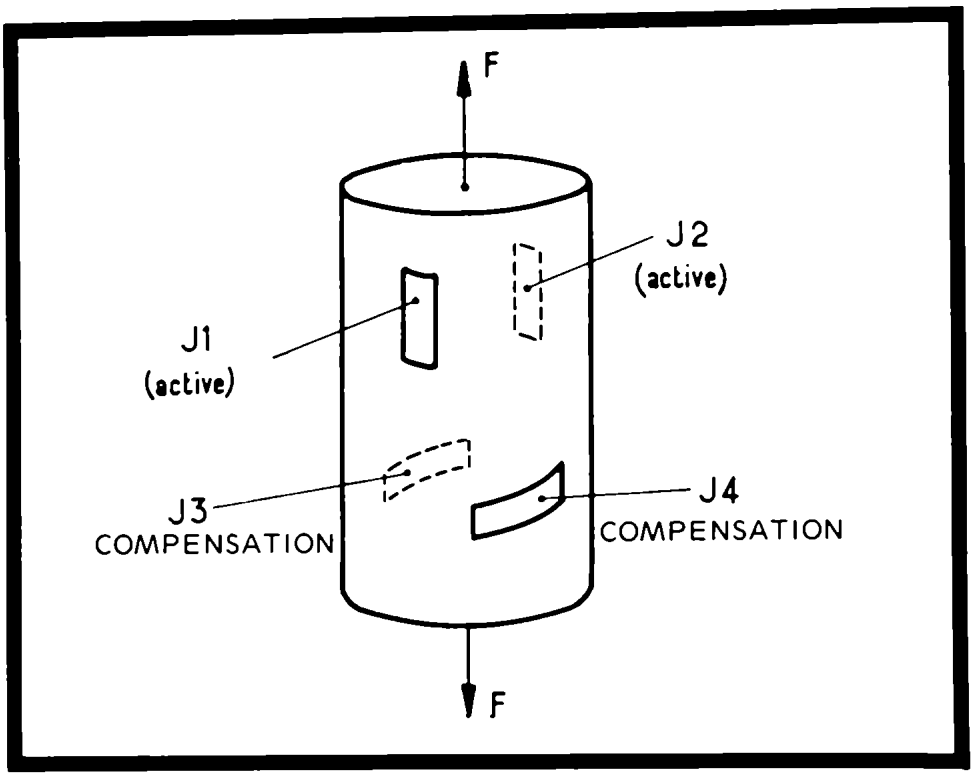


Fig. 2-9. Diagram showing the positions of gauges on a traction bar.

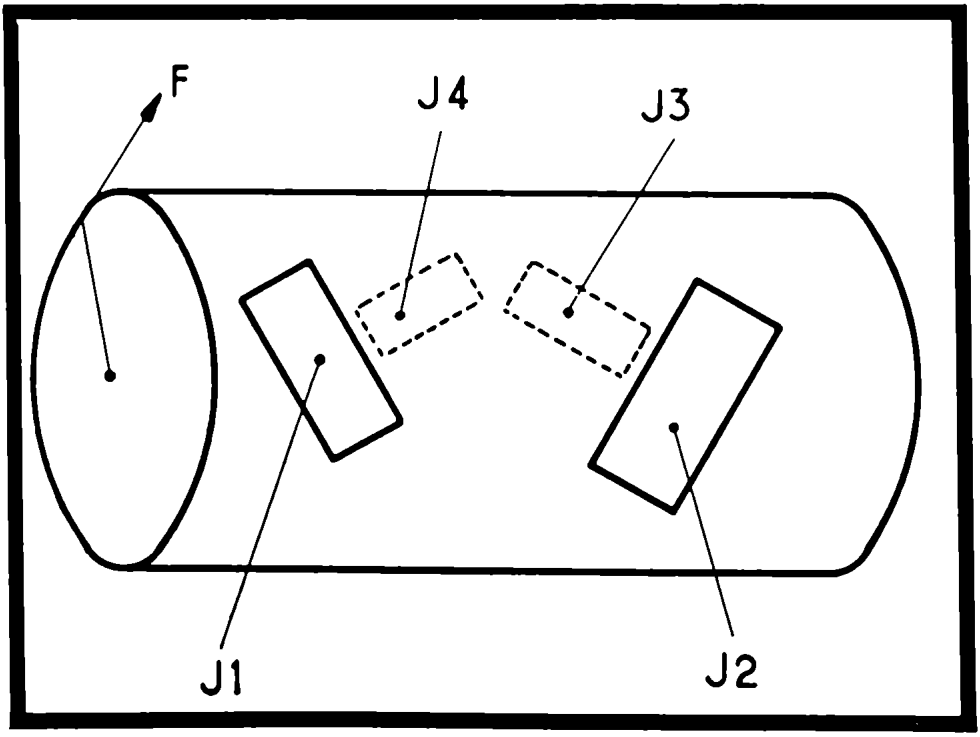


Fig. 2-10. Strain gauge positioning on a traction bar.

cells under the principal supporting points of a plane, knowing the empty and the fully-fueled weights, the actual weight can be measured, and the remaining fuel quantity can be deduced.

To measure the torque exerted on a cylindrical bar, four strain gauges (all active) are placed at 45 degrees, regarding the cylinder generatrix, corresponding to the highest dL over L of the fibers of the base material (Fig. 2-10). J1 and J4 (opposing J1 in the bridge circuit) are cemented at 45 degrees in the direction of the force F, and J2 and J3 are perpendicular to those fibers. The four bridge edge points are connected to insulated slip-rings connected solidly with the bar; contacts are made by fixed brushes. Contact variations occurring between slip-rings and brushes do not jeopardize accuracy because they are exterior to the proper bridge circuit.

Dynamic strain gauge measurements are concerned with the variation in terms of the time of the quantity examined; the transient behavior is of interest here, rather than accurate measurement of a permanent state. This behavior can be displayed permanently by a graphic recorder and in a short-lived way by an oscilloscope. While scopes are broad frequency-band devices, the roll-off frequency of graphic recorders, unfortunately, is rather low, say, from 100 to 2,500 Hz.

Straight-forward DC amplifiers of sufficient gain are drift plagued; carrier-current amplifiers are preferred for stability reasons, though they appear a bit more complicated. Since the highest modulation frequency is limited to one fourth of the

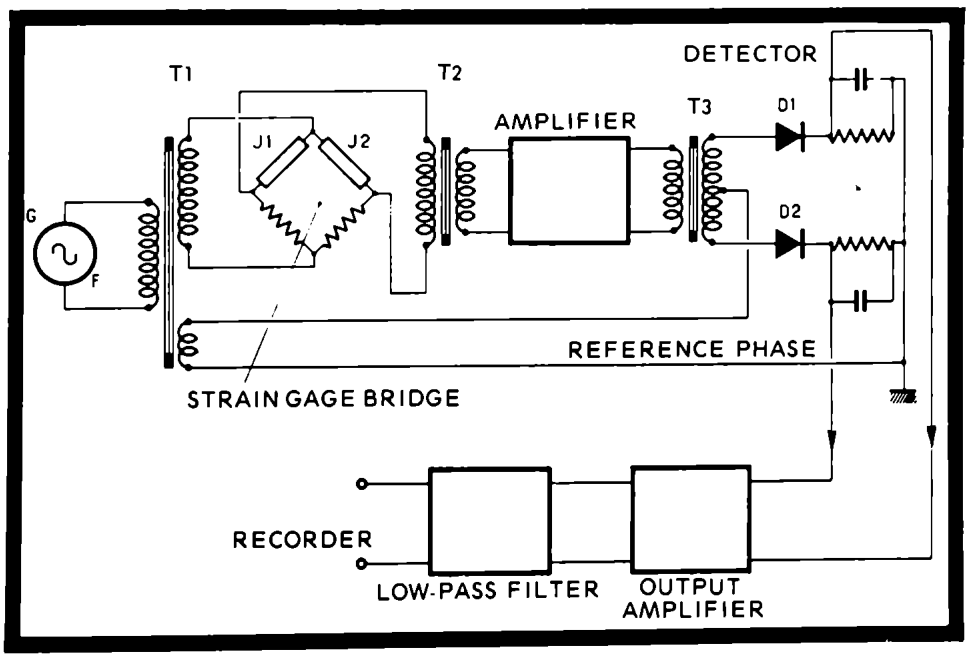


Fig. 2-11. Basic diagram of a carrier-current amplifier for extensometric measurements.

carrier frequency, the bridge supply (carrier) frequency generally is between 400 and 10,000 Hz, depending on the highest frequency component to be transmitted. There is no point in chosing too high a carrier frequency, for capacitive components become too worrisome.

A typical carrier current amplifier circuit is shown in Fig. 2-11. The bridge is supplied with carrier frequency F by transformer T1; its output, taken from the other diagonal, is applied by transformer T2 to a high-gain amplifier designed to pass F. Transformer T3, supplied by the amplifier output, gets its phase reference from T1 on the center tap and thus forms a phase detector with diodes D1 and D2. The resulting voltage is amplified by a low-gain DC amplifier with negligible drift. A low-pass filter eliminates the remaining carrier frequency component of the signal feeding the graphic recorder. It can be shown that the phase detector output is proportional to the resistive dR over R unbalance of the bridge and that the original polarity is restored, while a capacitive unbalance has no effect upon measurement accuracy.

Some years ago, semiconductor strain gauges composed of a silicon filament (microsensors) were marketed. They feature very interesting characteristics such as gauge factors (K) of 100 to 130, against about 2 for wire gauges. Furthermore, K is positive with p-type silicon and negative for n-type material. The high K, as well as the choice of polarity, allow for using a much lower gain amplifier, thus simplifying the overall circuitry. However, just like all other semiconductor devices, these gauges are very sensitive to temperature changes that not only vary the gauge resistance, but also the K factor and the length of the gauge. Of course, these variations can be compensated for, but the bridge then becomes more complicated. As a result, the opportunity of simplifying the amplifier is questionable if the resulting complexity is greater. However, for dynamic measurements not requiring a high accuracy, these silicon gauges may be interesting.

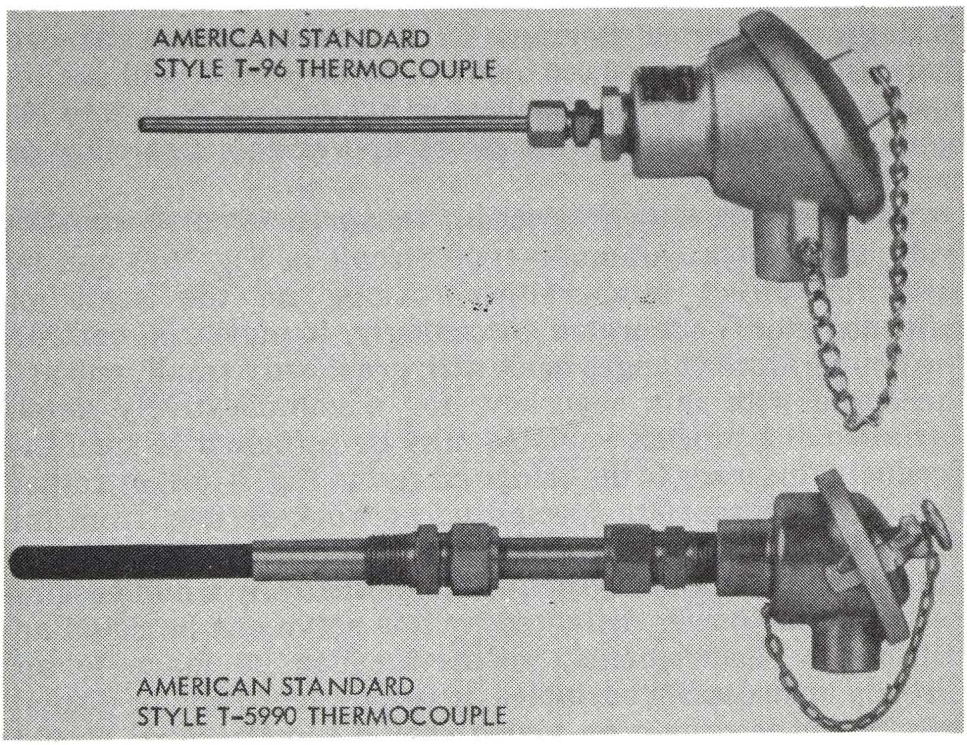
RESISTIVE TEMPERATURE TRANSDUCERS

The variation in resistance of a conducting material in terms of temperature can be used for (electrical) temperature measurement. In general practice, temperature-induced circuit variations are inconvenient, and to be minimized. They are emphasized here to make temperature measurement as easy and accurate as possible. Pure metals have higher (positive) temperature coefficients than alloys. One such copper-nickel alloy, constantan, owes its name to its rather temperature-independent resistance, making it a natural for

the manufacture of wirewound precision resistors. On the other hand, a copper wire undergoes a 2 percent increase of resistance for a temperature rise of 9 degrees F. For measurement of comparatively high temperatures, resistive gauges made of metals such as platinum are used.

Thermistors are composed of a bit of sintered metallic oxides with two connecting wires attached. This compound is very temperature-sensitive; it features a negative temperature coefficient an order of magnitude larger than that of a copper wire—a temperature rise of 1 degree F results in a decrease of about 2 percent of the device resistance, an easily workable variation. Thermistors are serviceable up to about 600 degrees F, and they are available in various forms such as washers, discs, rods and beads. Large cross-sections are required to pass comparatively large currents, allowing for direct control of a relay, for instance. A simple on-off temperature control for low-power furnaces using a thermistor is shown in Fig. 2-12. Similarly, a thermistor can be cemented on a machine part to monitor its temperature.

Because of their thermal inertia, large-size thermistors have a somewhat sluggish response to a quick temperature



Thermocouple designed for heat treating and other temperature-sensing applications. The T-96 may be used with temperatures up to 2,000 degrees F, the T-5990 up to 3,000 degrees F. Manufactured by Aero Research Instrument Div., American Standard, Inc.

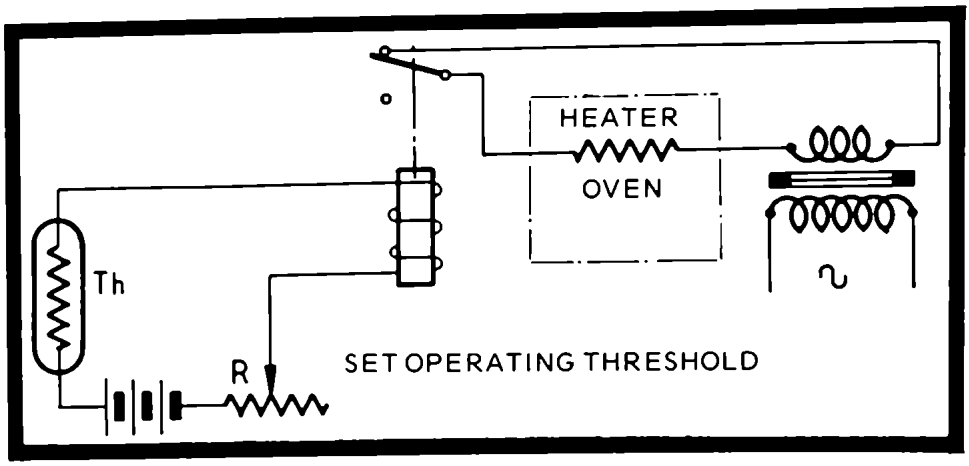


Fig. 2-12. Thermistor-controlled thermostat circuit for direct oven temperature regulation.

change, while the small bead type (about the size of a pinhead) reaches a thermal equilibrium in a matter of seconds. These tiny devices are available unmounted as well as glass-coated for chemical and biological mediums to be monitored.

Thermistor resistance is controllable in the manufacturing process and can be made quite low (a few ohms) to fairly high (100,000 ohms and more) referred to an ambient of 20 degrees C, or 68 degrees F. A typical high-resistance thermistor of about 120,000 ohms at 68 degrees F will see its resistance decrease to 4,600 ohms at 212 degrees F, and to 700 ohms at 356 degrees F. This offers appreciable flexibility to the designer.

If the temperature variation encountered is reasonably wide, the simple ohmmeter-type circuit in Fig. 2-13 can be used for direct temperature reading on the scale of galvanometer G calibrated accordingly. In industrial control, the temperature changes dealt with are mostly small, and are measured with fair accuracy, for a comparatively small variation of a process variable may be sufficient to spoil the product. This leads to the bridge circuit in Fig. 2-13B, or rather to a more appropriate specialized network derived from the basic Wheatstone bridge. Such a device must be balanced to get a reading; this can be done automatically by a null-seeking servo. As the power delivered by this servo can be rated according to requirements, the device can be made to act directly upon control variables, to print, record and-or display results and to trigger an alarm should something go wrong.

A flow of gas can be sensed using two bead-type thermistors, TH1 and TH2, both immerged in the gas (Fig. 2-14). However, while TH1 is placed amidst the gas stream, TH2 is exposed only to still gas. The source (B) not only powers the bridge, but also slightly heats by the same amount that the

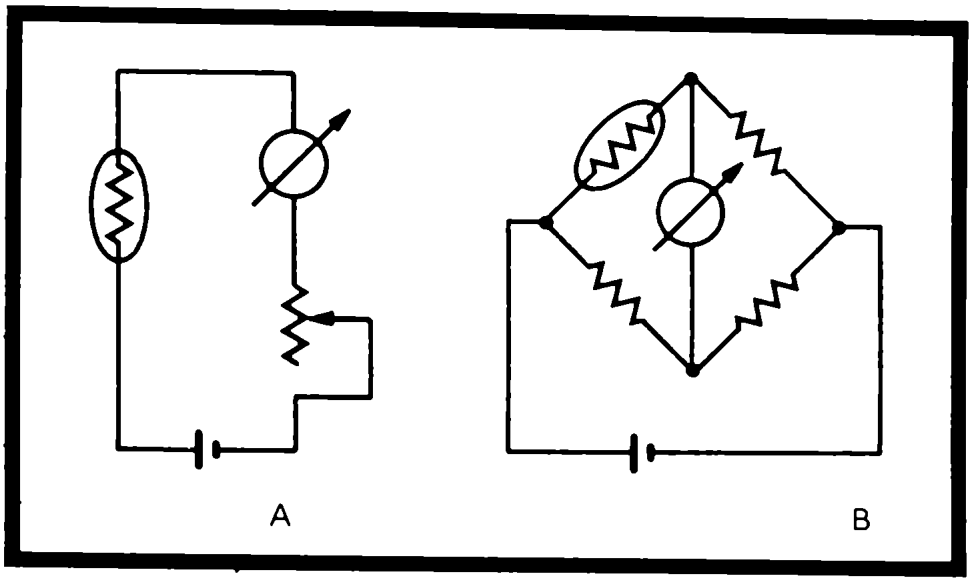


Fig. 2-13. Thermistor circuits for wide (A) and small-range (B) monitoring or measurement.

thermistor beads are conduction-cooled by the gas ambient. TH1 is convection-cooled by the gas stream, and the resistance ratio of TH1 and TH2 is a measure of the gas flow.

LIQUID CONDUCTIVITY AND pH MEASUREMENTS

Conductivity or resistivity, the opposite magnitude, indicates the acidity of an aqueous solution and, therefore, is an

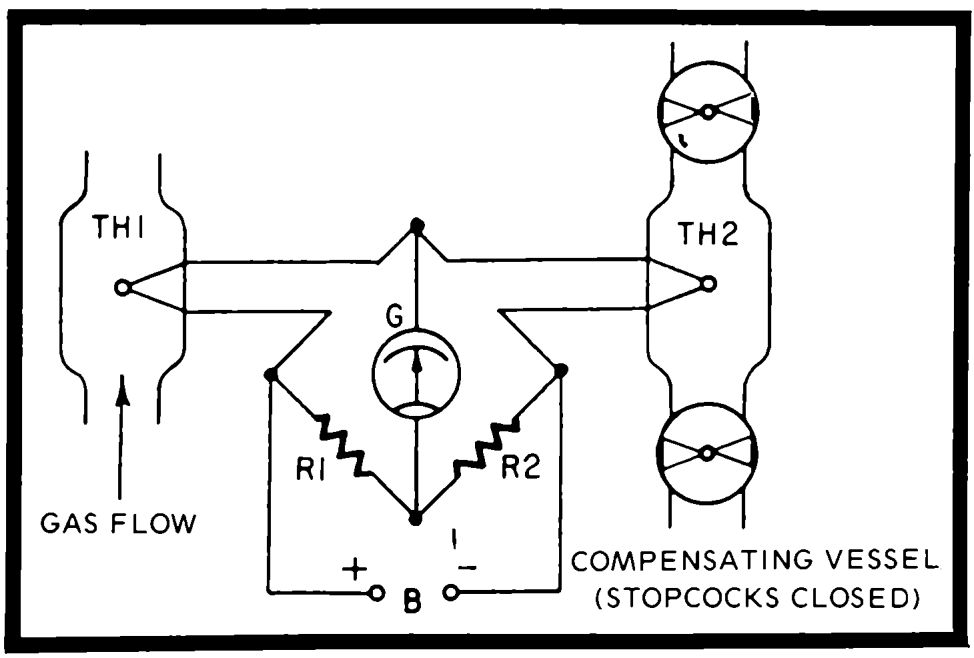


Fig. 2-14. Gas flow monitoring diagram using thermistors.

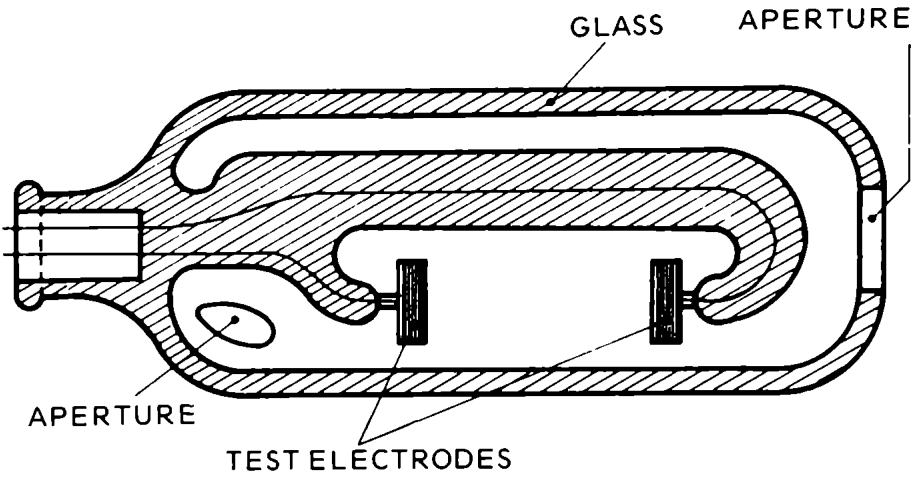


Fig. 2-15. Liquid conductivity monitoring probe.

important parameter in chemical processing. The conductivity cell in Fig. 2-15 is a generously apertured glass vessel containing two plane electrodes facing each other and brought out through insulating sleeves to a suitable connector. This device is immersed in the liquid to be monitored so that the electrodes are bathed by the liquid stream. The conductivity (or resistivity) of the medium comprised between the electrodes is measured by a bridge method, and valves are opened or closed in terms of the required variation of acidity. The bridge is AC rather than DC-operated to avoid polarization.

The acidity or alkalinity of an aqueous solution is measured by the pH; that is, the **negative logarithm of the H** (hydrogen) **ion concentration**, expressed in ions-grams per liter. While this scientific definition can be forgotten, it is good to remember that the pH scale extends from 0 (concentrated acid) to 14 (concentrated base), a solution being chemically neutral when the pH is 7 at 25 degrees C or 77 degrees F.

The pH plays an overwhelming role in the fluids that compose an important part of the living organism; only very small departures from the ideal value are tolerated. The normal pH of man is 7.4, with possible variations between 7.3 and 7.5; a pH factor outside these limits means death. This explains the major interest in pH measurement shown at least by the pharmaceutical, food, beverage and personal care products industries preparing stuff to be ingested by humans or to be applied to their skin.

The basic principle of pH measurement is illustrated by Fig. 2-16. A vessel divided by a porous membrane contains in one partition the solution of an unknown H ion concentration, and in the other a solution of a known constant ion concentration. A measurement electrode and a reference electrode are dipped in these solutions; the whole device forms a galvanic cell whose electromotive force, expressed in mv at 20 degrees C (68 degrees F), is given by:

$$E = K - 58\ (pH)$$

where K is a constant, depending on the type of electrodes used. E varies linearily with the pH; a unity pH change always corresponds to a 58 mv change of E, whatever electrode types are used.

Electronicians will be interested to know how such a pH cell devises suitable measurement techniques. This cell offers a very high resistance of up to 100,000 meg and even more. Furthermore, even a very small measuring current polarizes the cell; i.e., produces a reverse electromotive force at the electrodes, making the device useless. For the cell to remain unaltered, the measuring current is not to exceed 10^{-11} to 10^{-12} amperes, requiring use of an electrometer tube or field effect transistor. The electromotive force to be measured is in the millivolt order of magnitude. The output voltage of the electrometer or FET used as an impedance converter is applied to

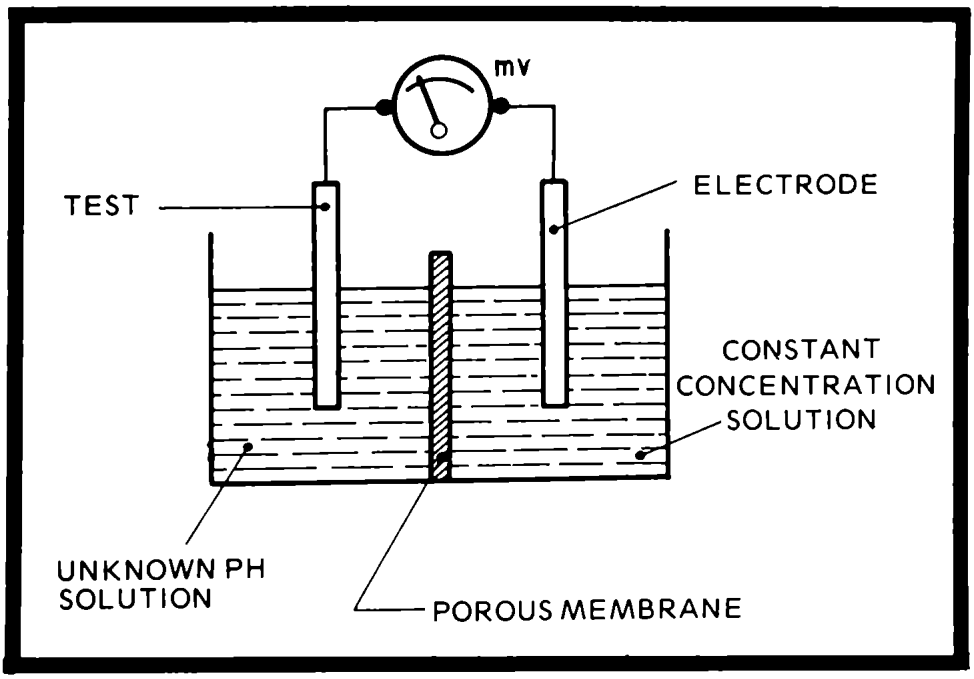


Fig. 2-16. Sketch illustrating the basic principle of pH measurement.

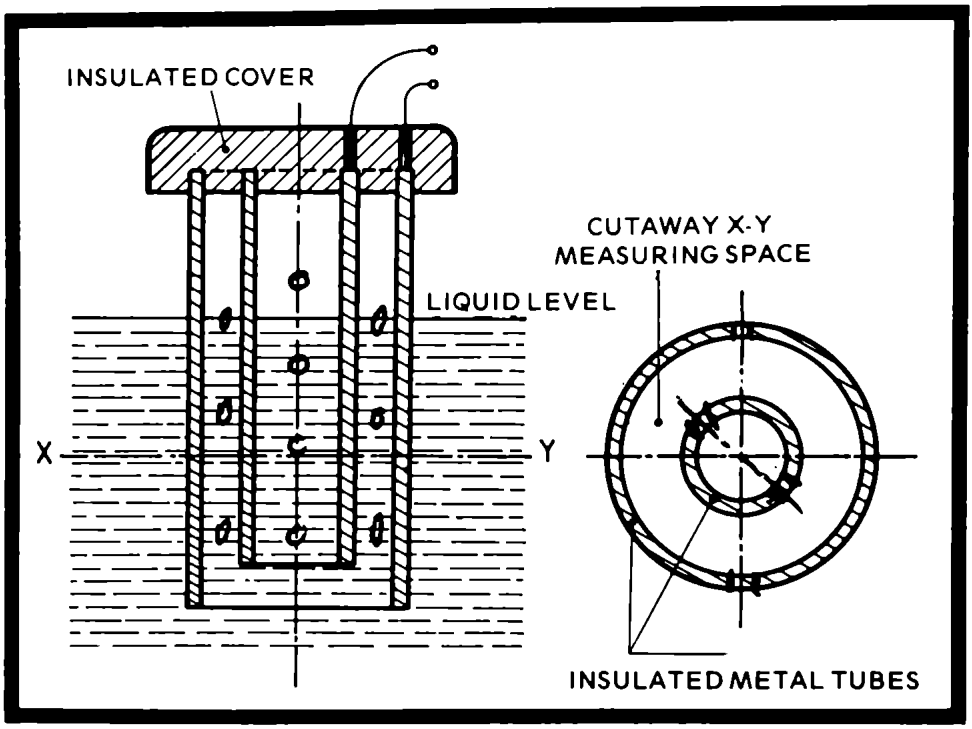


Fig. 2-17. Capacitive transducer using the dielectric method of liquid-level monitoring.

a null-seeking potentiometer, and the measured pH value is recorded or printed out and possibly displayed by a galvanometer pointer or, more recently, digitally.

The electrode problem is rather complex and cannot be treated in a few words. Practically, the glass electrode and the calomel electrode are extensively used. These electrodes already contain the membrane and the constant ion concentration solution; it is necessary only to dip such an electrode in the solution to start the pH measurement process.

CAPACITIVE TRANSDUCERS

The capacitance of a capacitor composed of two parallel plates is proportional to the surface (of one plate) and to the dielectric constant of the medium existing between the plates, but inversely proportional to the gap length separating the plates. Capacitive transducer operation is based, mostly, upon the variation of the gap length or displacement of the plate surfaces facing each other by the magnitude to be controlled.

Though uncommon, the dielectric constant variation is quite workable in some specific cases, such as, for instance, a liquid-level transducer for dielectric liquids such as gasoline. As shown in Fig. 2-17, this wholly static device is composed of

two mutually insulated coaxial assembled metal tubes extending from the tank top down to the bottom. As the dielectric constant of air is K equals 1, that of the liquid is somewhat higher (1.5 to about 3). The capacity of this coaxial capacitor increases with increasing liquid level, and the quantity of liquid contained in the tank can be deduced from a capacity measurement.

Such a transducer (and others to come) has a rather high capacitive reactance (Zc) at low frequencies, owing to:

$$Z_c = \frac{1}{2\pi fc}$$

f is the measuring frequency and C the transducer capacity. Indeed, impedance as well as resistance measurements are delicate and inaccurate for very low as well as for very high values. Therefore, it is a good plan to work the bridge at higher audio frequencies, and even at F. But even then, the error voltage cannot be readily transmitted to a distant point. In the case of 4-capacitor bridges, such as shown in Fig. 2-18, the high-impedance output is applied to the gates of Q1 and Q2 (FETs) used as an impedance converter, with the output linked to a (possibly remote) amplifier, followed by a phase discriminator supplied with a phase reference (by the bridge oscillator powering the bridge) via a separate twisted and shielded cable.

A typical high-frequency circuit for working capacitive transducers is represented schematically in Fig. 2-19. Transistor Q1 oscillates at a frequency, f, mainly determined by inductance L and transducer (variable) capacitance C. The oscillator voltage is amplified by buffer amplifier Q2 and the output is applied to a frequency discriminator circuit tuned to mid-frequency f. The discriminator output characteristic represented below gives the discriminator output voltage V_0 in terms of a capacity variation centered on V_0 equals 0, ideally comprising fairly linear segments of both polarities, symmetrically, with regard to the zero crossing.

Notice the successive transformations of the measured magnitude as required by the measuring process: a **liquid level** is converted into a **capacitance** which tunes an oscillator in which the **frequency** is discriminated to produce a **voltage** that an amplifier can transform into a **force** to act upon something. Many control systems use a number of such parameter conversions.

In most capacitive transducers, the variation of the magnitude to be sensed is converted into a change in the geometry of the transducer capacitor. The pressure of a gas in a pipe can thus be sensed by means of a metallic bellows B

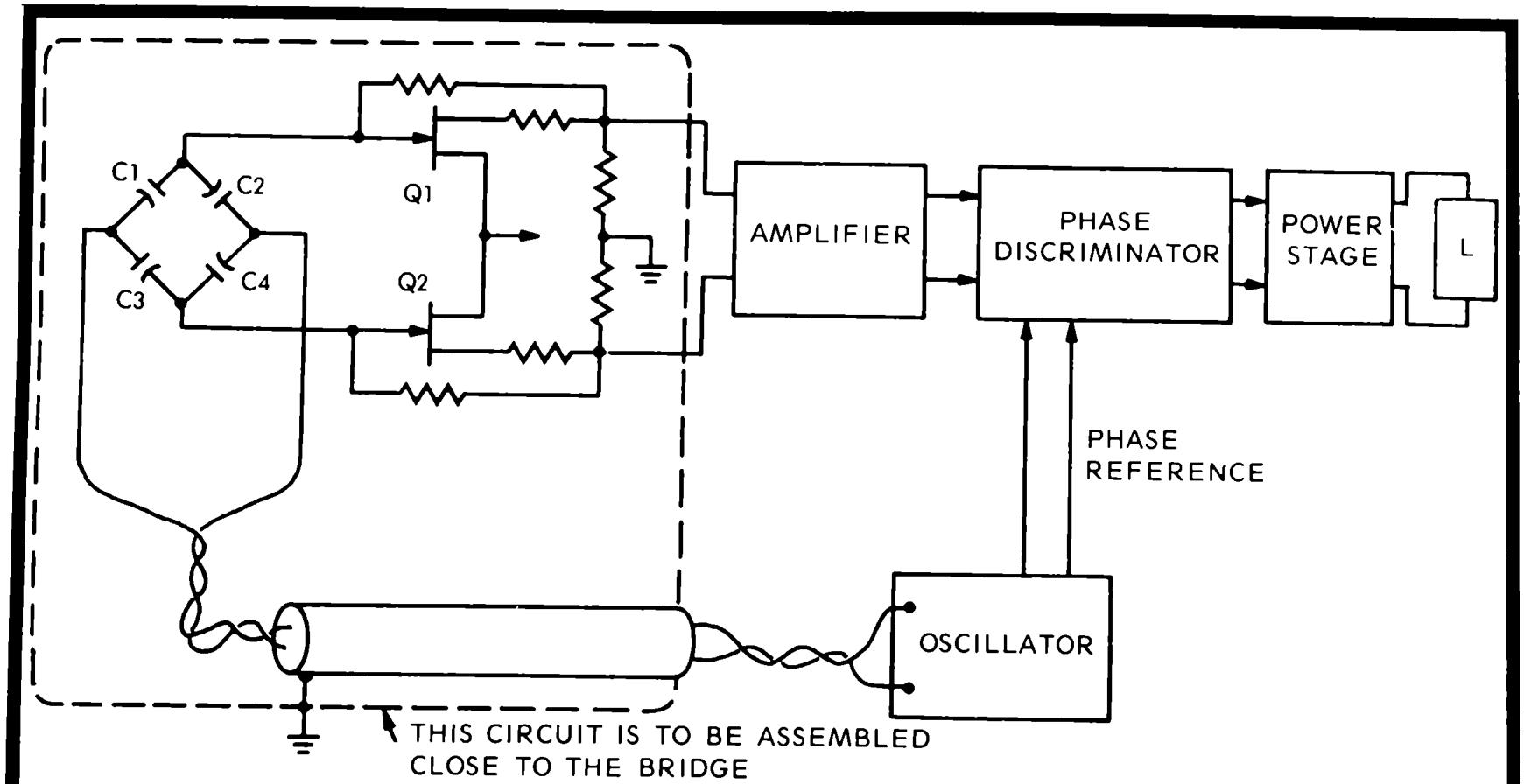
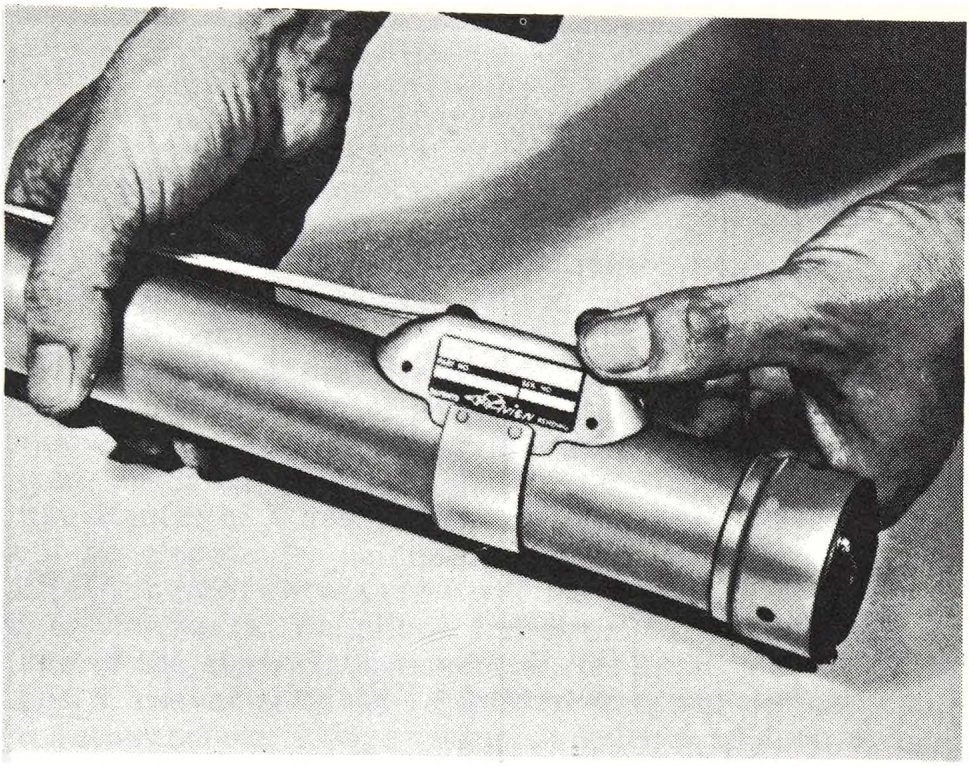
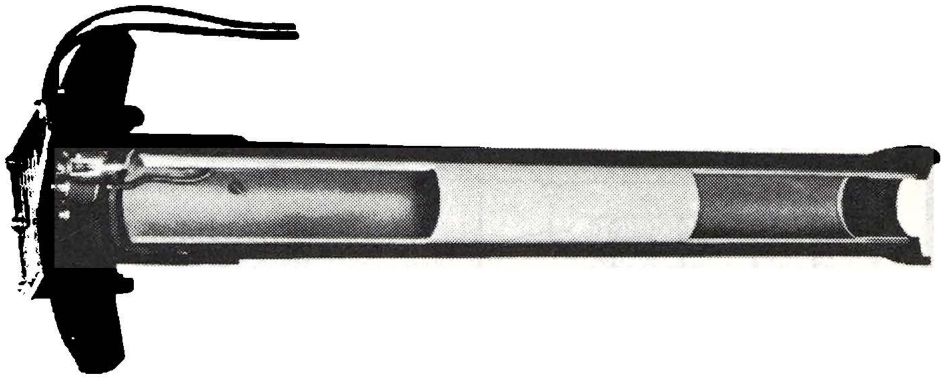


Fig. 2-18. Schematic of an FET impedance converter for high-impedance coupling.





The liquid is the dielectric in this capacitive liquid level transducer. Cutaway view (bottom) shows the inner and outer capacitor plates. As the liquid rises, it displaces air as the dielectric, thus increasing the capacitance across the transducer terminals. To compensate for variations in the dielectric constant across the transducer terminals. To compensate for variations in the dielectric constant with temperature, a thermistor can be added as shown in the top photo.

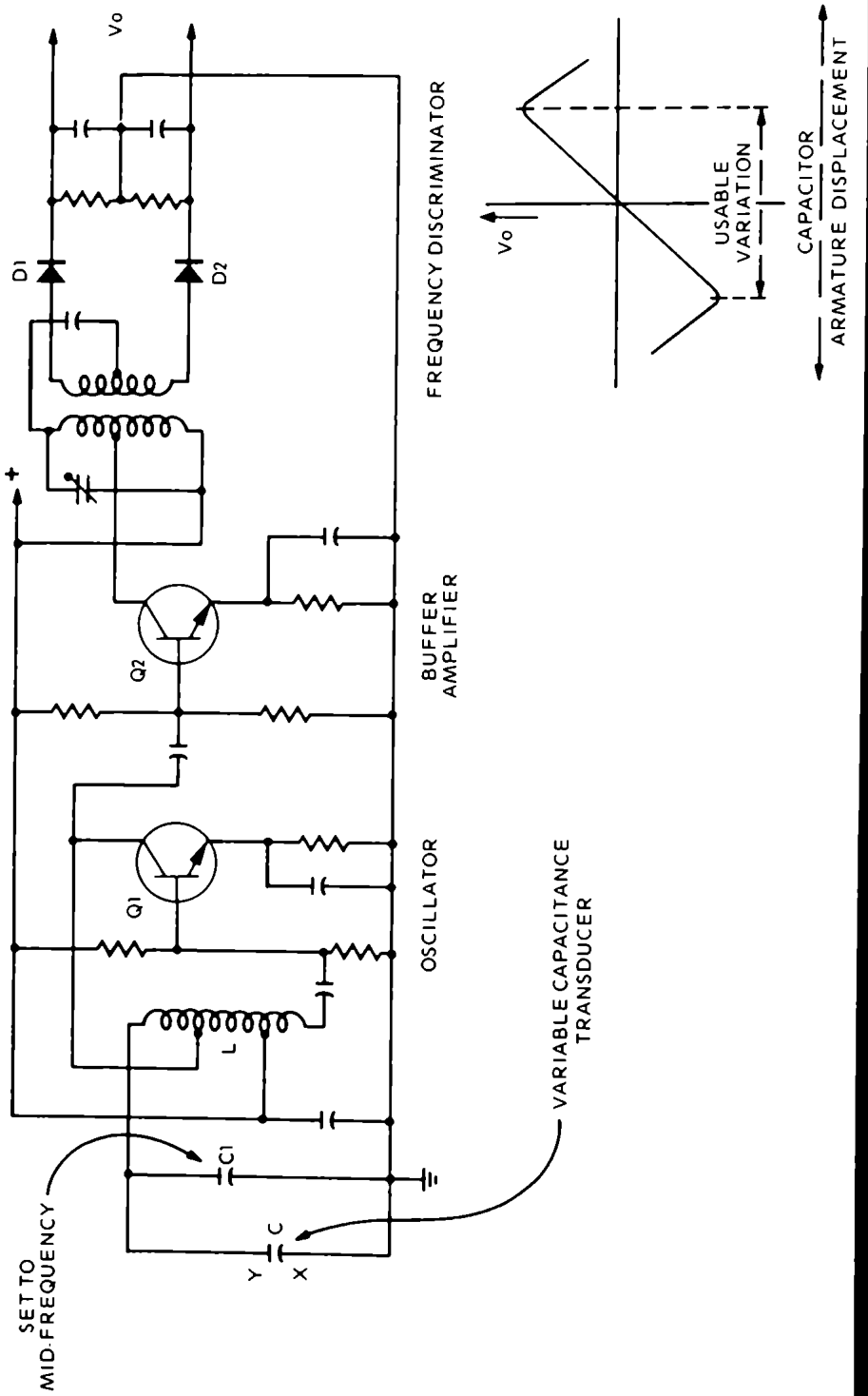


Fig. 2-19. Capacitive transducer circuit designed for RF applications.

connected to pipe P. A pressure increase will distend B and thereby bring (earthed) armature X nearer to fixed insulated armature Y, lowering the capacitance of the transducer (Fig. 2-20). The accuracy of such a transducer will be jeopardized, however, by any factor other than pipe pressure which varies the length of the gap between X and Y.

This shortcoming can be eliminated by a differential transducer, because the quantity to be measured now is the ratio of two like magnitudes, rather than the absolute value of this magnitude. In the differential transducer in Fig. 2-21, the angular displacement of a shaft is sensed by a (grounded) vane (W) connected solidly to the shaft facing more or less fixed insulated plates X and Y. Calling x and y the capacitances of X and Y, respectively, to ground, the device will be electrically balanced if x equals y; any angular displacement of the vane now will increase x while decreasing y, or vice versa. Magnitude and direction of this unbalance (and of the shaft displacement by the same token) can be sensed conveniently by a bridge supplied by an audio frequency source. Bridge output is applied to an error voltage amplifier followed by a phase discriminator to restore the direction of unbalance.

The thickness guage in Fig. 2-22 applies to any type of material, although this is still another type of differential capacitive transducer. The inner (earthed) armature of a 3-plate capacitor is connected solidly with the test spike resting on the body of which the thickness is to be measured. If this

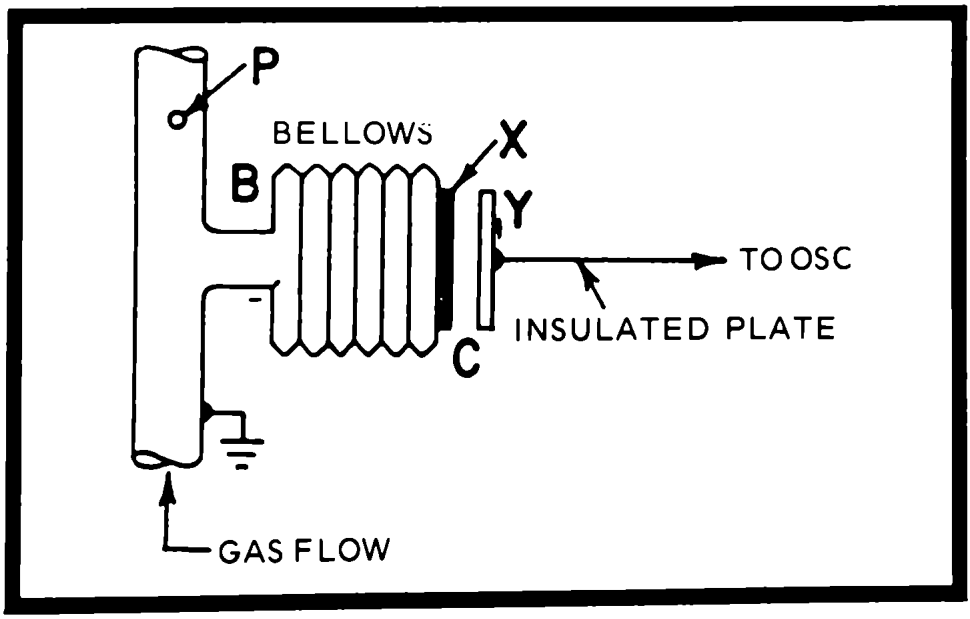


Fig. 2-20. Drawing of a capacitive pressure transducer.

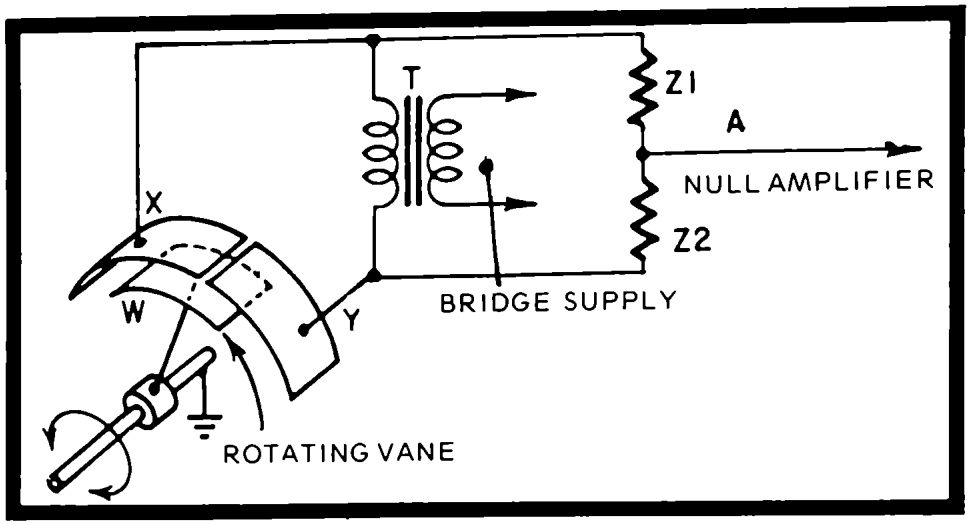


Fig. 2-21. Differential capacitive angular displacement transducer.

body is somewhat thin or small, the inner armature will lower, thereby increasing the lower and decreasing the upper element capacitance, which are measured as formerly.

The simple two-armature transducer, however, is quite workable for a dynamic plot of the compression ratio, or pressure-volume relationship of internal combustion engines. For easy insertion in the combustion chamber, the device is shaped like a sparkplug, but its bottom is covered by a thin

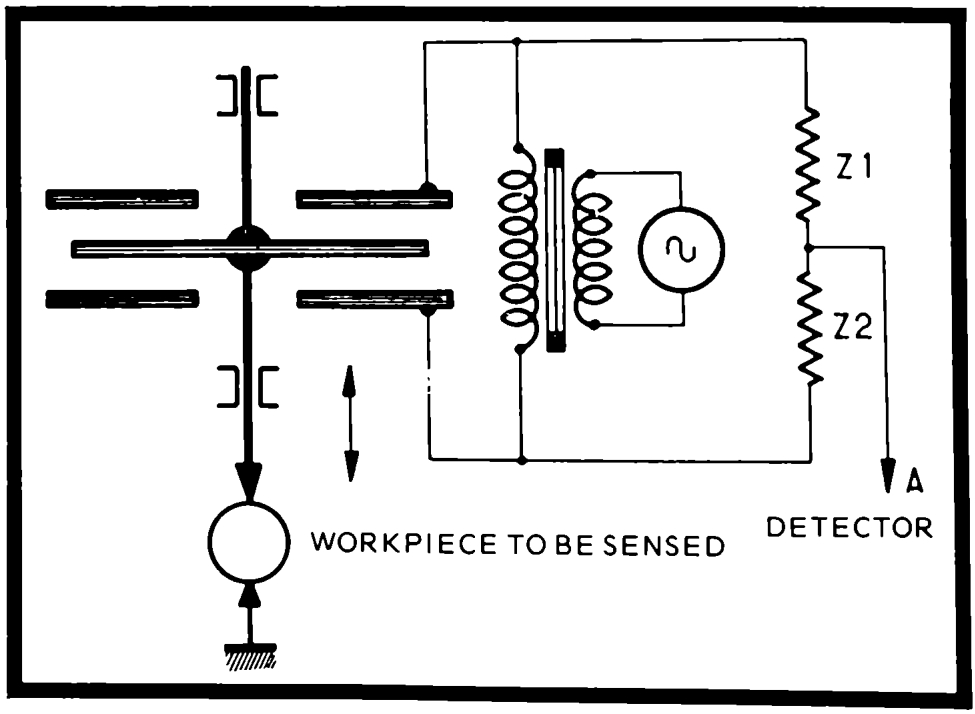


Fig. 2-22. Differential transducer for small linear displacements.

metallic membrane (armature X) facing an insulated fixed electrode, Y (Fig. 2-23). Pressure changes deform the membrane and vary the gap between X and Y, hence the transducer capacitance. This magnitude, however, is small and its variations are still smaller. The best way to emphasize the variations may be by tuning an oscillator and applying its output to a frequency discriminator.

INDUCTIVE TRANSDUCERS

In a **variable reluctance transducer**, the displacement of armature A of an iron-core inductance results in a reluctance change, hence a variation of inductance detected by an AC bridge (Fig. 2-24). Since several other parameters can result, too, in a misleading inductance change, the differential

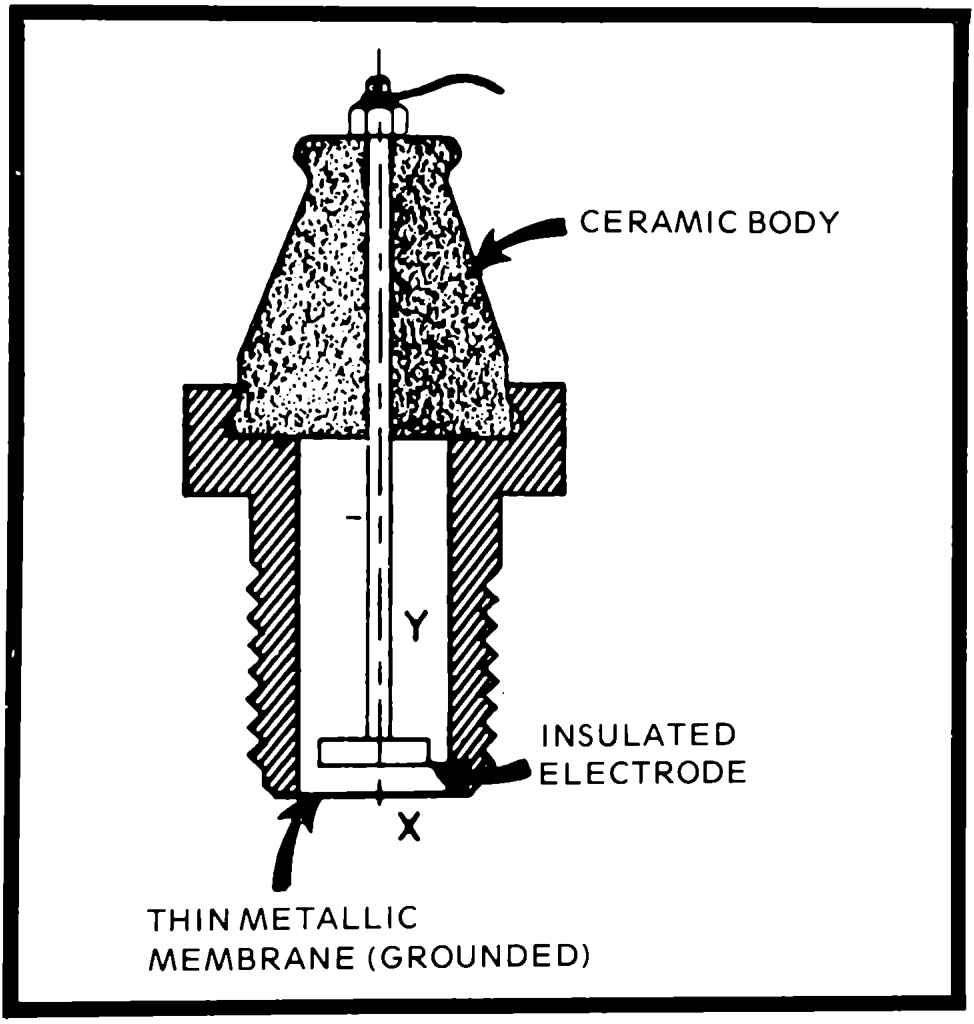


Fig. 2-23. Capacitive transducer for the dynamic plot of pressure-volume diagrams of internal combustion engines.

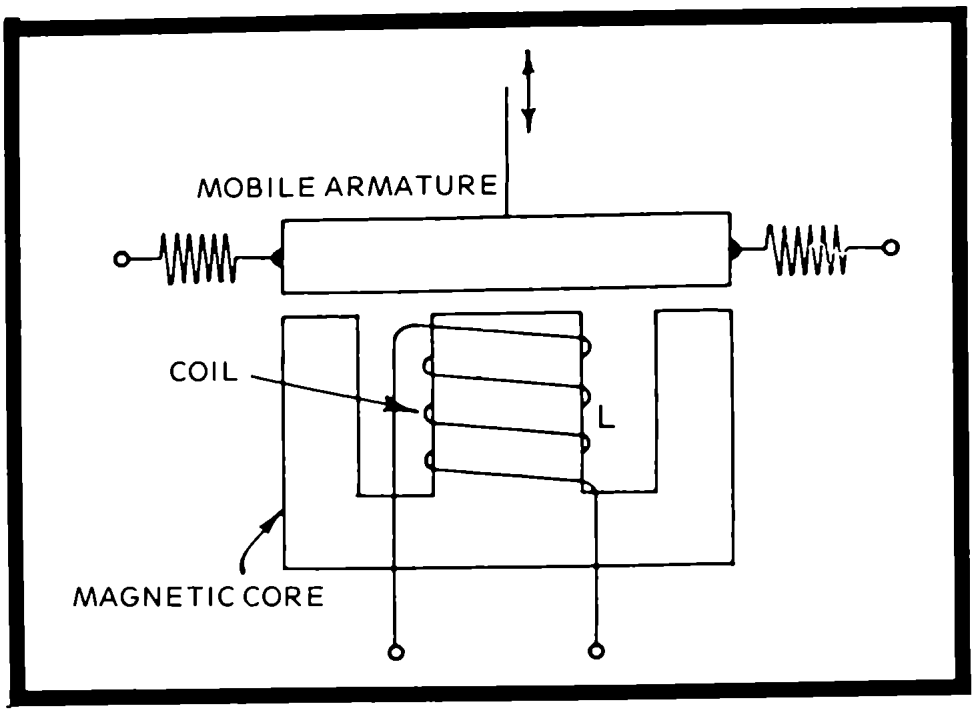


Fig. 2-24. Variable-reluctance linear displacement transducer.

operation mode is to be preferred here because such perturbations may readily cancel. Furthermore, the impedance range of inductive transducers is well suited for bridge circuitry, and such a transducer already contains two bridge (ratio) arms.

In the differential variable-reluctance transducer in Fig. 2-25, armature A, solidly connected with the object whose displacement is to be sensed, is introduced between two facing iron cores supporting like coils, L1 and L2. The displacement of A is such that the air gap of one core is reduced while the other is increased, resulting in opposite variations of L1 and L2. The bridge is completed by two like impedances, Z1 and Z2, it is AC-fed by a transformer. The rectifier meter shown will display only the unbalance, not its actual unbalanced direction; the direction can be obtained by means of a phase reference using a phase discriminator.

The accelerometer in Fig. 2-26 is an application of this principle. Acceleration of a plane, a rocket, car or other moving object is evidenced by mass M flexibly supported within the vehicle. An armature A bolted to M is inserted between two facing iron cores solidly connected with the outer structure and supporting coils L1 and L2. If the vehicle accelerates, the inertia of mass M opposes the change of pace, and a rearward movement (i.e., opposite to the direction of acceleration) results. Contrariwise, the mass tends to push

forward if the moving body slows down, just as the passengers of a suddenly braked car are thrown forward. The displacement of A regarding the cores is detected as formerly. Notice that this transducer is sensitive only to horizontally-directed accelerations and to the horizontal component of a

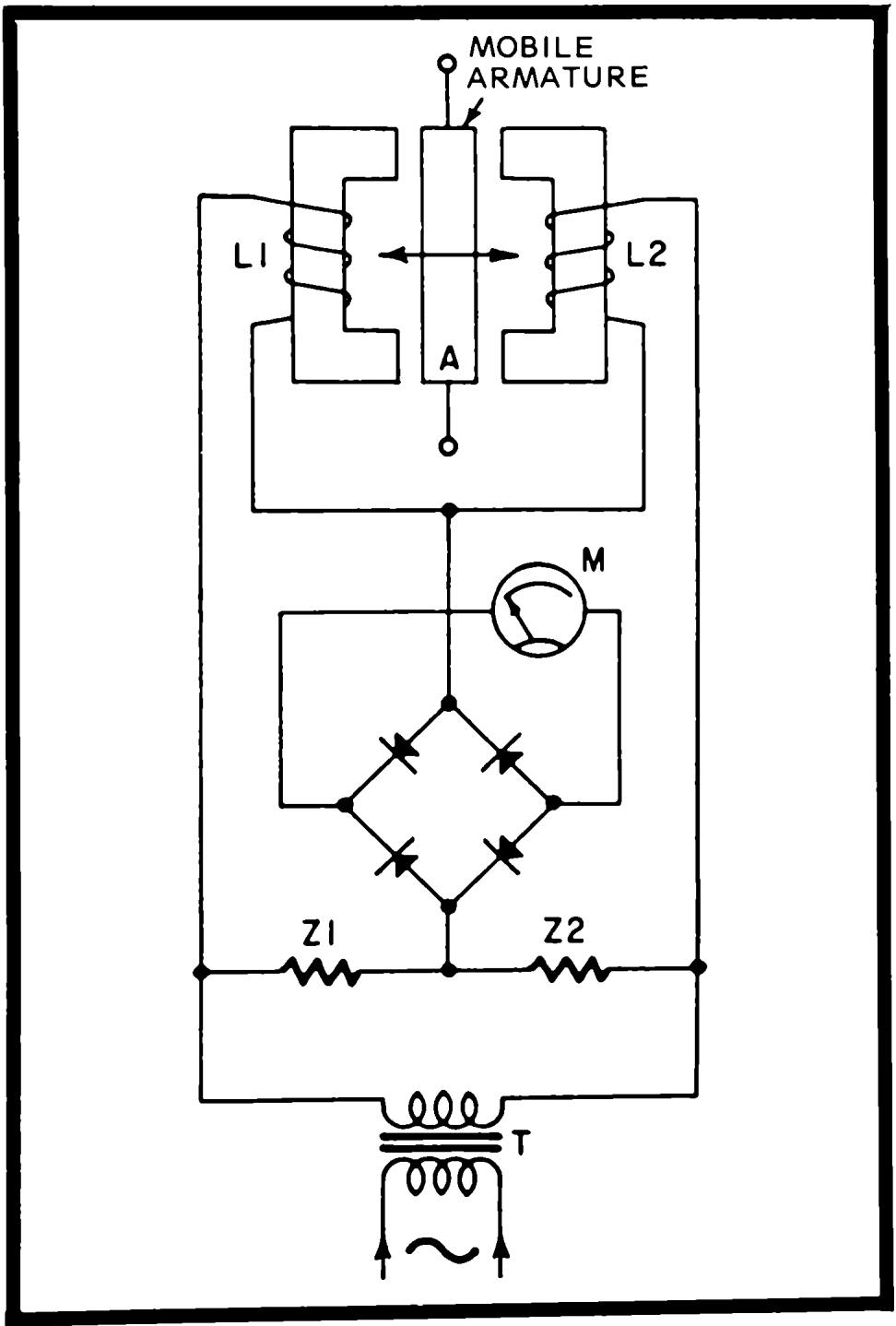


Fig. 2-25. Diagram of a differential variable-reluctance transducer.

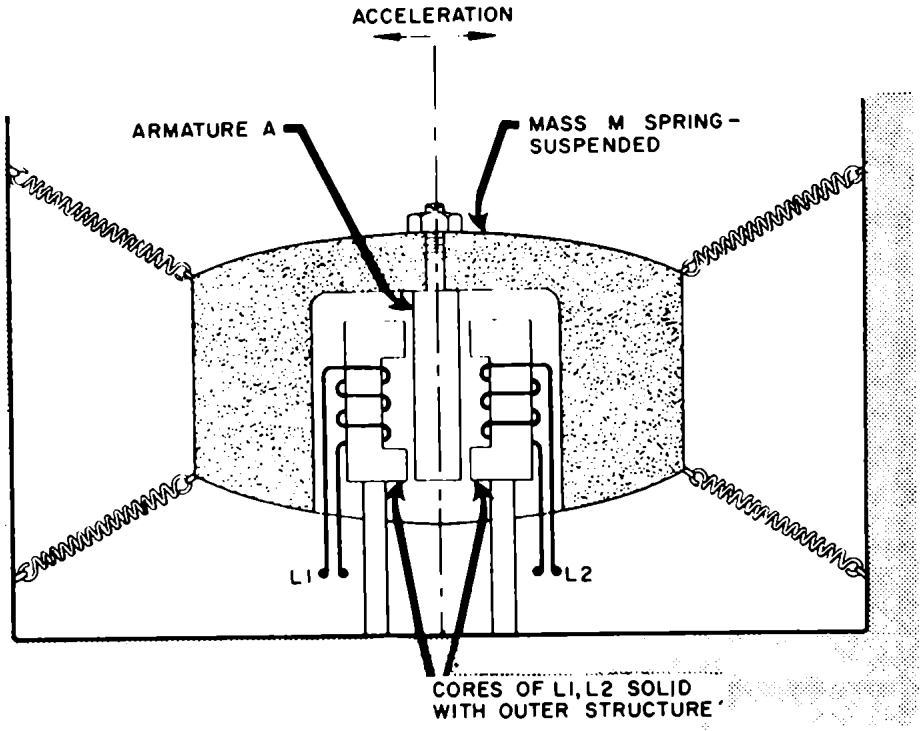


Fig. 2-26. Accelerometer using a variable-reluctance differential transducer.

sloping acceleration vector; for orthogonal directions, additional transducers are provided. This allows for neatly separating efforts along plane or space coordinates. Similar devices are used to detect and analyze machine and structure vibrations.

Mutual-inductance transducers employ a different operating principle. The differential transformer in Fig. 2-27 is composed of a tube made of a nonmagnetic material supporting two like coils, L1 and L3, series-connected in a way that their magnetic fields are in opposition. A third coil, L2 is wound between L1 and L3. A slug made of a magnetic material can glide along the coil axis; it is linked to the body whose displacement is to be sensed.

Suppose, now, that L2 is fed an AC voltage. Provided the slug is magnetically centered, equal voltages are induced in L1 and L3, and as they oppose, the resulting output voltage is zero. A slight axial slug displacement is enough to increase one induced voltage while reducing the other. The unbalance is evidenced by an output voltage in which the amplitude and phase contain the slug displacement amplitude and direction information which is extracted by a phase detector. Alternately, coils L1 and L3 can be AC-fed, and L2 will provide the output.

The differential transformer in Fig. 2-28 is composed of two C-shaped iron cores supporting AC-fed primary windings L3 and L4, as well as secondaries L1 and L2 which are series-connected so the induced voltages oppose each other. A slug made of magnetic material is linked to the body whose displacement is to be sensed; it moves along the common gap axis, closing one gap while leaving the other. When the slug is magnetically centered, no output appears across Terminals a and b. A slight upward movement of the slug increases the voltage across L1, while reducing that across L2, so that a net difference voltage develops between a and b.

Fig. 2-28 also shows a **ring modulator** that detects displacement amplitude as well as direction by means of a phase reference supplied by the supply line to one modulator diagonal using a center-tapped transformer secondary. A zero-center DC milliammeter can be calibrated for a direct displacement amplitude reading, or, alternately, the modulator output can be applied to a suitable amplifier for action.

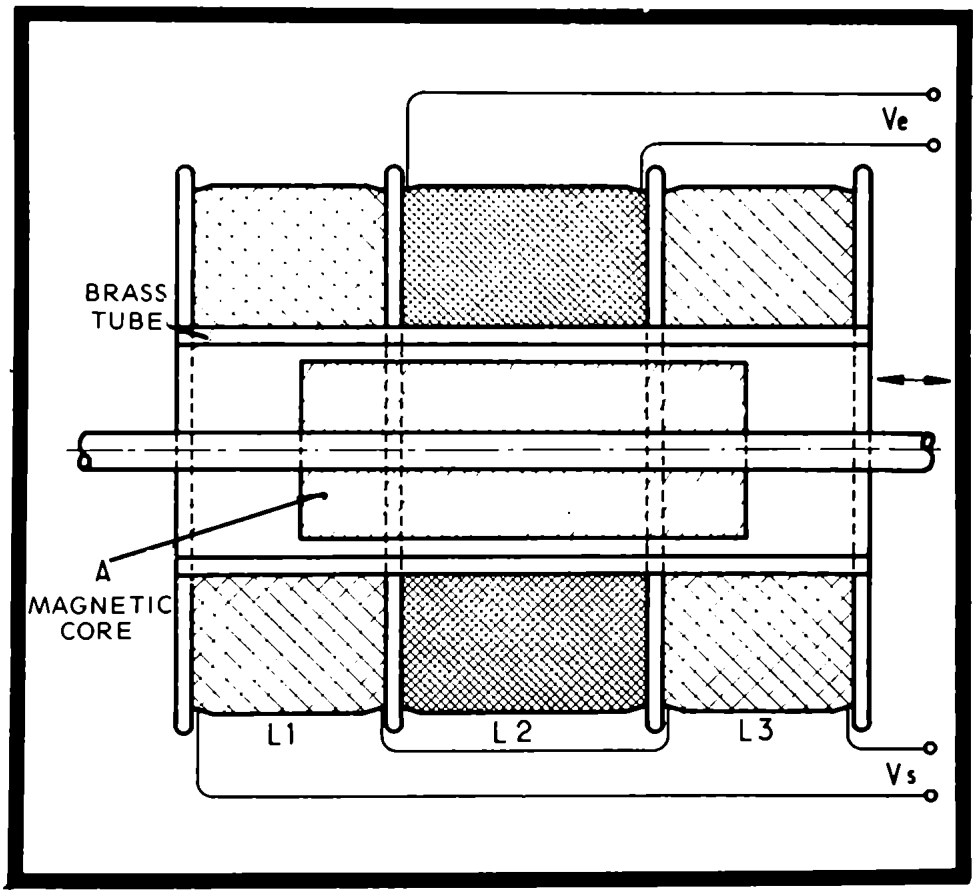


Fig. 2-27. Diagram of a differential displacement transducer.

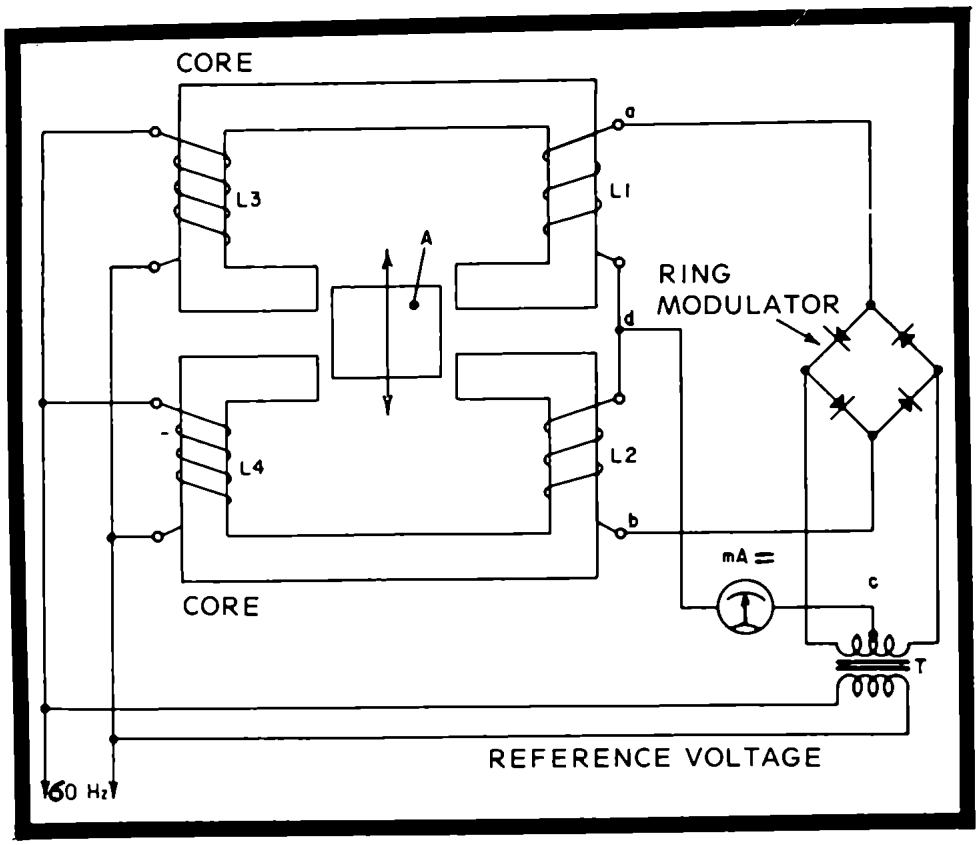


Fig. 2-28. Differential transformer connection diagram for small displacement sensing.

Unlike most other transducers, mutual-inductance types yield a directly usable output power level at a convenient medium-range impedance. Hence, amplifiers can sometimes be dispensed with, and small, low-power motors for null-seeking devices can occasionally be powered directly by the modulator output. These transducers, however, exhibit a nonlinear output-displacement characteristic, making them suitable for the detection of movements near the balance point rather than for linear displacement measurement. Differential transformers are extensively used in machine tool positioning controls.

ANGULAR POSITION TRANSMITTING SYNCHROS

A synchro is a special type of small power AC machine comprising a stator or fixed part supporting three poles spaced 120 degrees apart, and a rotor with a unique winding connected to two slip-rings. The stator coils of two such synchros are tied together as shown in Fig. 2-29, and rotor windings R1 and R2 are fed by an AC source. This electrical arrangement behaves like a mechanical link between these

rotors, and if we rotate the armature of the so-called "transmitter" synchro by a certain amount, the voltages induced in its stator windings will set up a magnetic field in the "receiver" synchro and rotate its armature by a like amount.

This is the basic principle used to monitor the remote angular position transmission of a lathe carriage spindle or a parabolic radar antenna. If the opposing torque of the receiver rotor can be neglected (no power being delivered), the angular position can be reproduced with an accuracy of possibly 0.1 to 0.5 degree. By means of a gear train without backlash, the angular rotation of the transmitter shaft can be multiplied. As a result, the information accuracy is increased accordingly. Except for the mechanical damper featured by most receivers that suppress system oscillations in the vicinity of the final position, both machines are alike.

In the **synchro-comparator** (Fig. 2-30), the AC-fed rotor, R1 is given, manually or automatically, an angular position to be occupied by the device under control; for instance, a radar parabola turret. The rotor of machine 2 (called synchro-comparator) is driven by the device under control, thereby reflecting its actual position. This rotor is not connected to a supply source; rather, it delivers an error voltage fed to the input of a power amplifier; the amplifier output powers the turret driving motor, possibly by means of a rotary amplifier. If the angular position of R1 is varied, the stator field of 2 induces an error voltage across R2. Once amplified, this signal

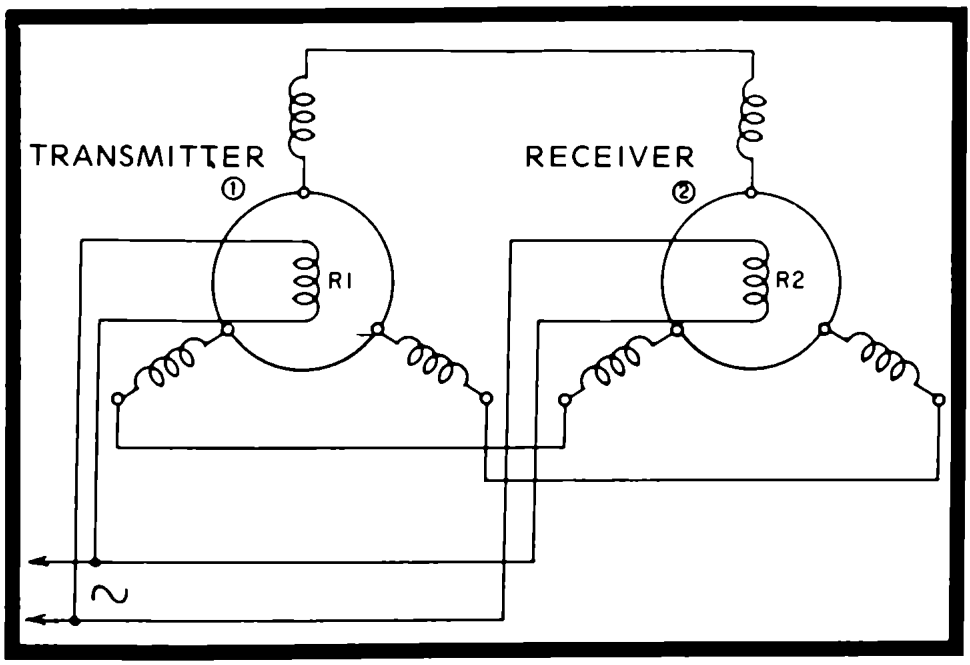


Fig. 2-29. Schematic of an angular position transmitting and receiver system.

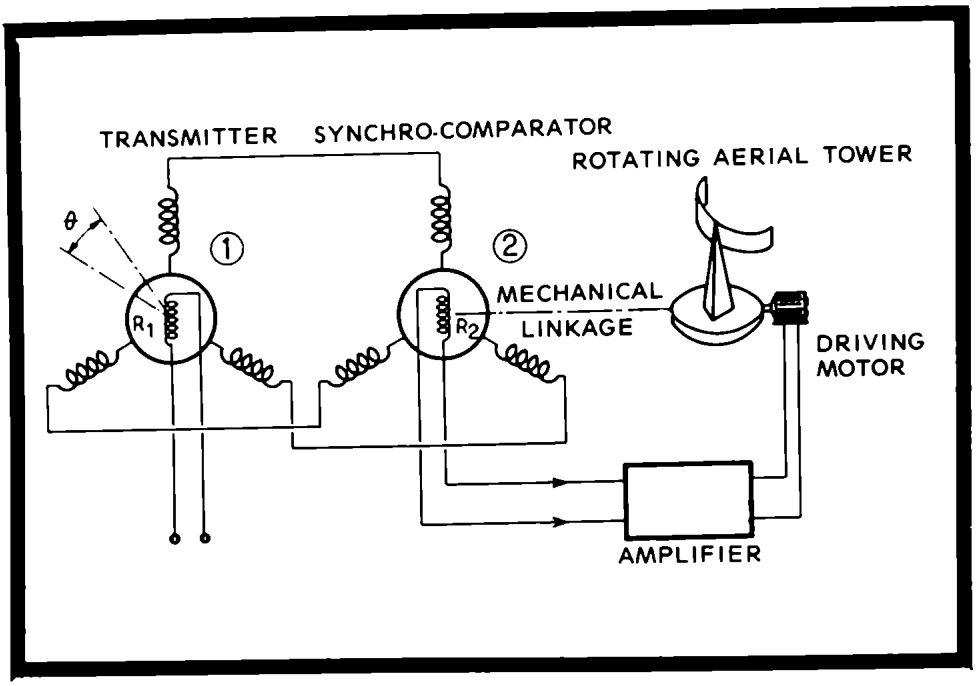


Fig. 2-30. This synchro-comparator system conforms the angular position of a rotating system with the position of rotor R1.

powers the driving motor to rotate the turret as well as R2 in a way to cancel the angular difference between both the rotors. As soon as this is done, the system comes to a stop.

The **differential receiver** is another rotary synchro comprising three stator and also three rotor windings connected to the 3-pole stators of transmitters 1 and 2 as shown in Fig. 2-31. As long as R1 and R2 occupy the same angular position, nothing happens, but if one of these armatures is rotated by an angle θ, the stator field of the differential machine rotates by the same angle regarding its rotor field (or vice versa), and if R3 is left free, it will rotate by the same angle θ. This actually is an error detector displaying the sum or the difference of the angular positions of transmitters 1 and 2. This system is used, for instance, to display the synchronizing efficiency of a system comprising a number of power drives in a steel rolling mill. In multi-jet planes, differential receivers are used to add the shaft revolutions of the individual fuel counters in order to totalize the quantity of fuel burned to determine how much of it is left.

In this system, rotor R2 of transmitter 2 can be disconnected from the AC supply and tied to an amplifier input, and the output of the same amplifier, in turn, powers the driving motor. The transmitter thus becomes a synchro-comparator. This arrangement is called a **differential transmitter** (Fig. 2-32). Rotor R2 is mechanically driven by the system under

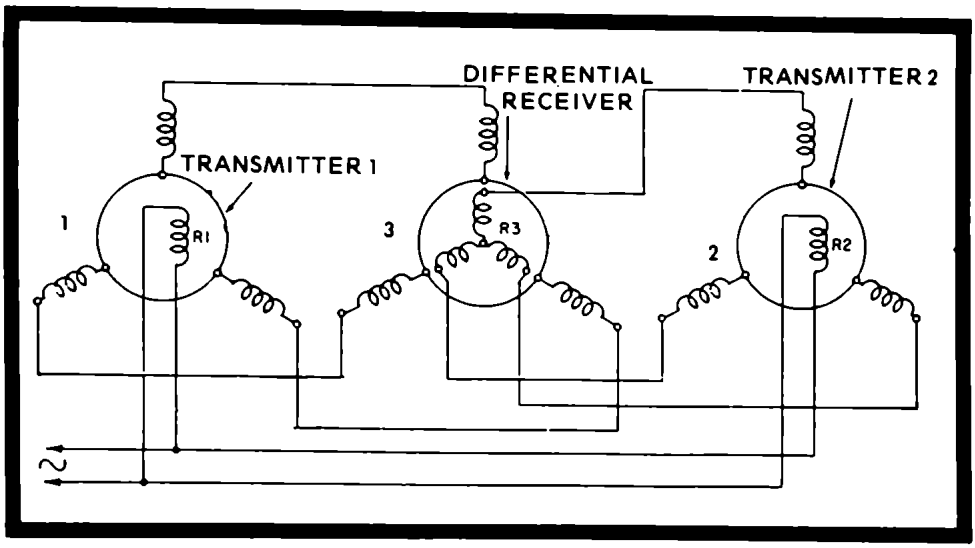


Fig. 2-31. This differential receiver displays the sum or difference of two angular positions, R1 and R2.

control, be it the work table of a milling machine whose angular position θ_2 it reproduces. Rotor R2 is given an angular displacement θ_1, and after another rotation, θ_3, a correction is applied to R3. The error voltage, ξ, collected on the slip-rings of R3 in this case, is given by:

$$\xi = k\ [\theta_2 - (\theta_1 + \theta_3)]$$

k is a constant. Hence, the system calculates the angular difference between its proper position (identical with that of the device under control) and the sum of two read-in positions.

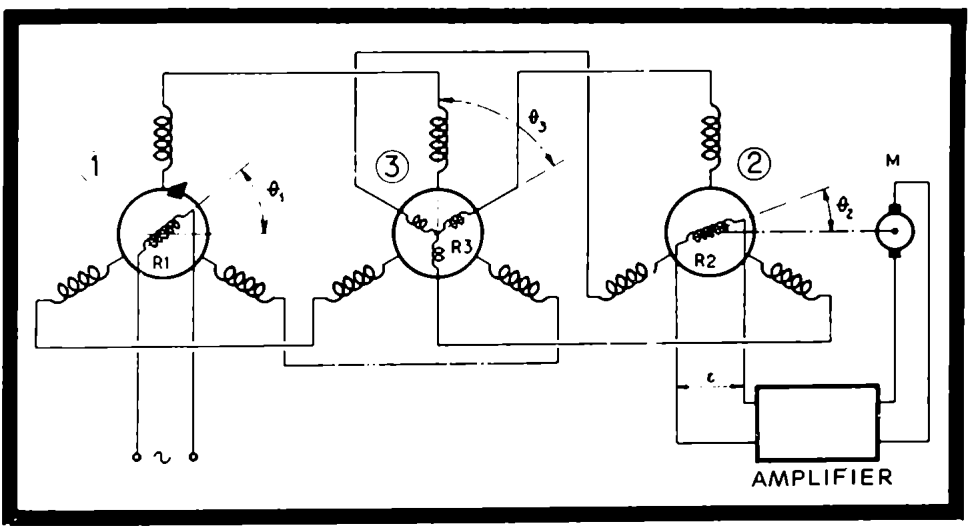


Fig. 2-32. Schematic of a differential transmitter which calculates the angular difference between its proper position and the sum of two read-in values.

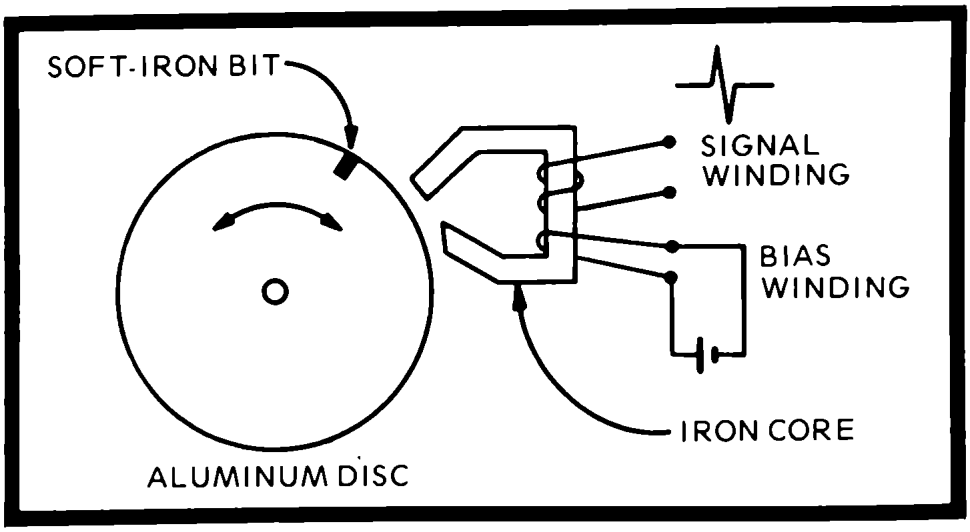


Fig. 2-33. An inductive rotation transducer of the type sketched produces one pulse per turn.

PULSERS FOR REVOLUTION COUNTING

The various transducers examined above are suited for sensing a limited angular variation; they cannot monitor continuously rotating machine shafts to totalize the number of revolutions. If this count is carried out during a unit time, a minute for instance, the number of turns counted equals the number of turns per minute.

We will not consider here ohmic-contact devices because of their inherent unreliability and unavoidable wear. Furthermore, many small-power drives simply lack the power required to close a contact without slowing down. Therefore, only non-contact-making, preferably pulse-generating rotation transducers, are discussed.

An inductive rotation transducer that generates a pulse which is easy to detect and to count at every revolution is represented in Fig. 2-33. This is a pickup device similar to a magnetic recording head. It is composed of a magnetic core made of soft steel laminations or ferrite and contains a signal winding and a bias winding (the latter can be replaced by a biasing permanent magnet somewhere in the magnetic circuit). The pickup faces a wheel mounted solidly on the shaft whose rotation is to be monitored, leaving a small air gap between them. The wheel is made of a nonmagnetic material, with a small soft iron bit engaged in its outer surface. When the wheel rotates, the gap of the pickup core briefly narrows down once per revolution. When the rotating iron bit passes under the gap, a pulse of one polarity is induced in the pickup, followed by a pulse of the opposite polarity as the bit leaves the gap.

Using a photoelectric device (as explained later), it is possible to establish rather simple contactless rotation transducers by arranging a polished flat area on the shaft so that a light beam is reflected on a suitable light-sensitive device at a given angle (Fig. 2-34); for all other positions of the flat, only a diffused light of comparatively low brightness reaches the radiation receiver. This results in a neat pulse per revolution. Alternatively, the image of a brightly lit shaft on a dark background can be projected on the photoelectric cell. The illumination can be suddenly extinguished by blackening a shaft segment; in this case, a pulse of opposite polarity will still result.

TACHOMETER GENERATORS

This is another type of rotation transducer which generates a voltage closely proportional to the speed of rotation. A DC **tachometer** generator is a small DC generator designed to feature a fairly linear voltage-speed characteristic. It could be used to display the speed of a rotating machine on the scale of a D'Arsonval meter calibrated accordingly, or by some numerical readout. While this goal can be achieved by the rotation transducers described above, tachometer generators are extensively used for speed regulation in closed-loop control systems.

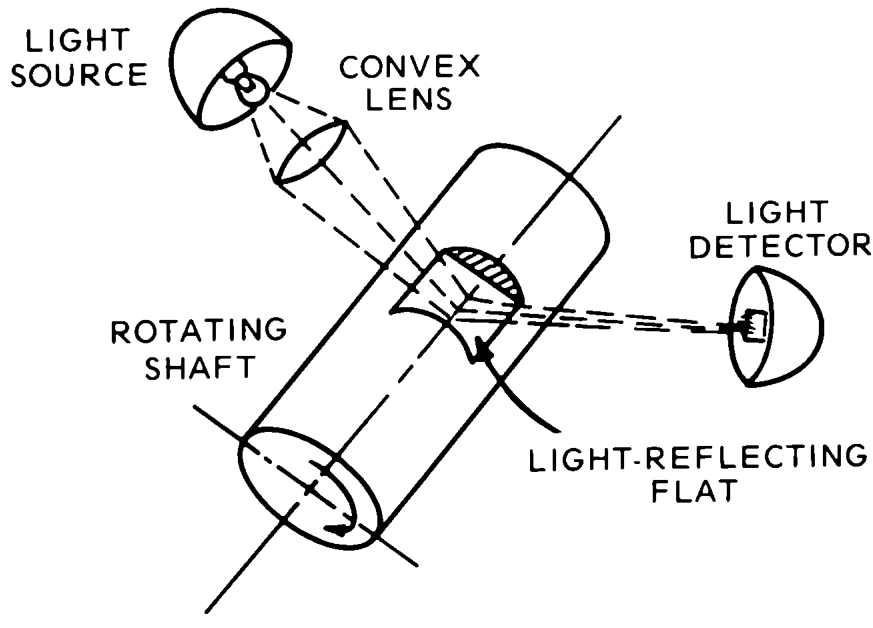


Fig. 2-34. Drawing of a photoelectric transducer.

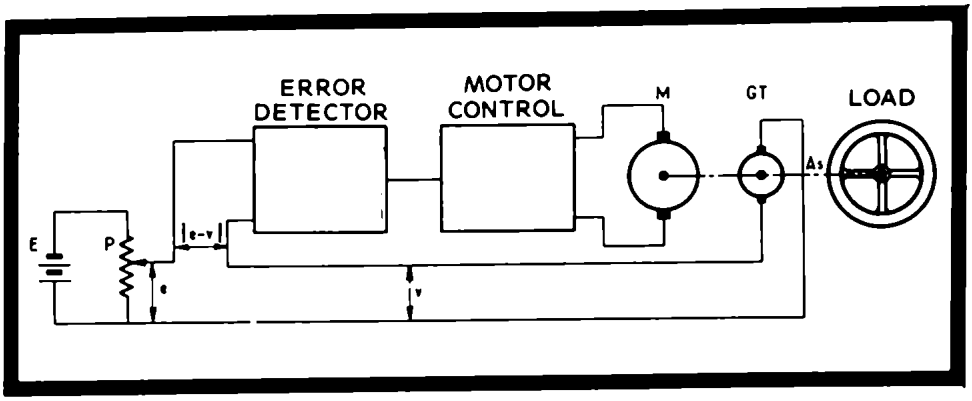


Fig. 2-35. Block diagram of a tachometer-generator speed control regulator.

The speed variation regulation of a DC driving motor is schematically represented in Fig. 2-35. The tachometer generator (TG) is generally assembled so it overchanges the end of the shaft of motor M. Its speed-proportional voltage v is compared to an adjustable DC reference voltage e by an error detector. The detector input is the difference e and v. This error detector instructs the thyristors, ignitrons or magnetic amplifier contained in the motor control to increase or to decrease the power supplied to the motor, thus accelerating or slowing down the motor according to the error signal polarity. With the error signal nulled, the motor rotates at a constant speed. A load increase will tend to slow down the motor, but the change is immediately counteracted by a decrease of v, and the resulting error voltage acts to compensate for the load change. The rate speed is determined by e and it can be set by potentiometer P, which acts as a motor speed governor. Resetting P immediately fixes a new working speed.

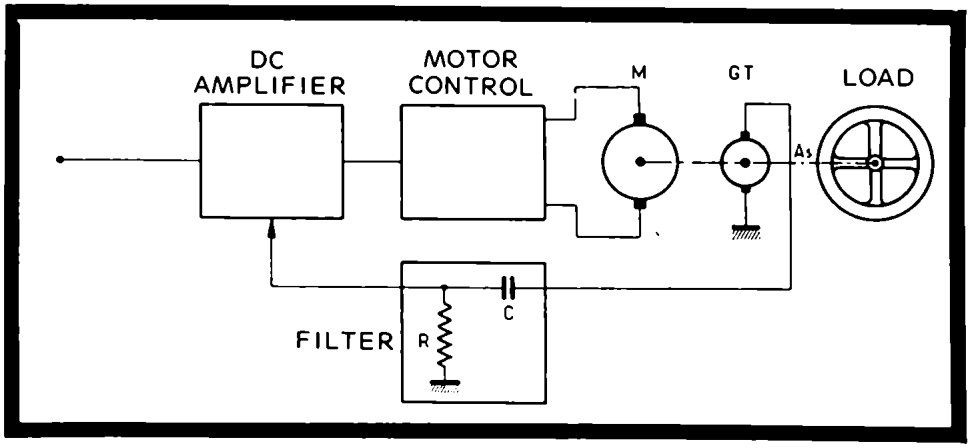


Fig. 2-36. Tachometer-generator block diagram showing method used to correct the response characteristic of a closed-loop system.

A different type tachometer generator application is shown in Fig. 2-36; it represents the final stages of a control system composed of a DC amplifier, a motor control and DC motor M1. Tachometer generator TG, sitting on the end of the motor shaft still produces a DC voltage, v, taken on its commutator brushes to be fed into the amplifier input as an inverse feedback. A series-connected RC high-pass filter removes the DC and slowly varying components; only rapidly varying voltages are transmitted. (This is a differentiating circuit.)

Hence, if the motor shaft is at a standstill or running at a constant speed, no inverse-feedback signal is transmitted by the filter, and the amplifier gain is high. If a speed variation occurs, and especially if the system tends to become unstable, capacitor C transmits the voltage variation, and, thereby, reduces amplifier gain. This results in reduced overshoot and positioning error and in a lower control delay.

AC tachometer generator operation is obtained by feeding one of the windings of a two-phase induction motor with an AC voltage of frequency (Fig. 2-37). Armature rotation induces in the free winding an AC voltage v of frequency f. The amplitude of the induced voltage is nearly proportional to the speed of rotation. If the motor reverses its direction, the phase of v reverses, too; an interesting way to use an AC tachometer in a carrier-current control system.

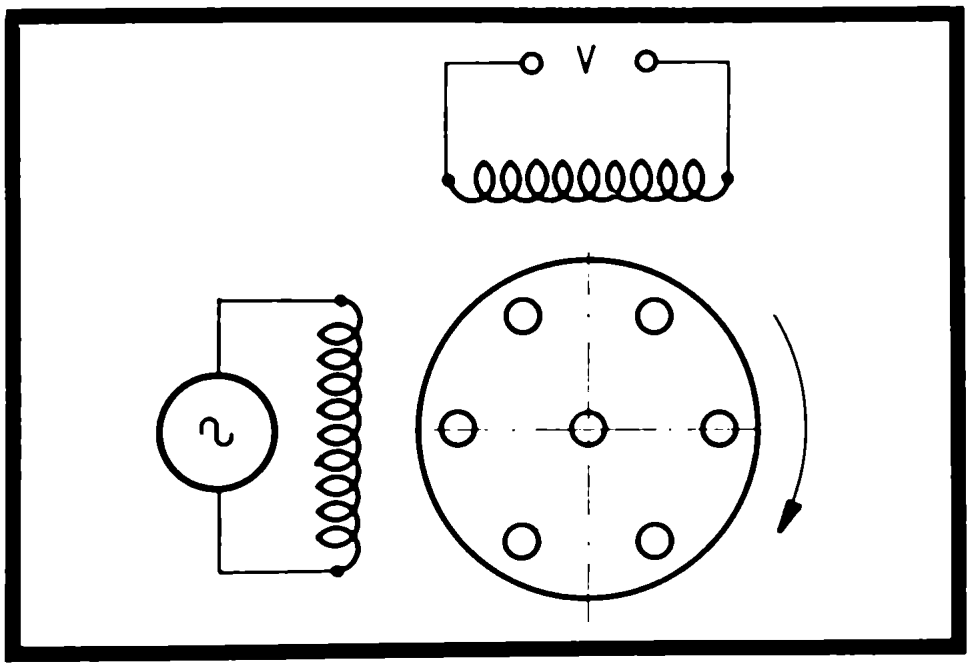


Fig. 2-37. A 2-phase induction motor is serviceable as an AC tachometer generator.

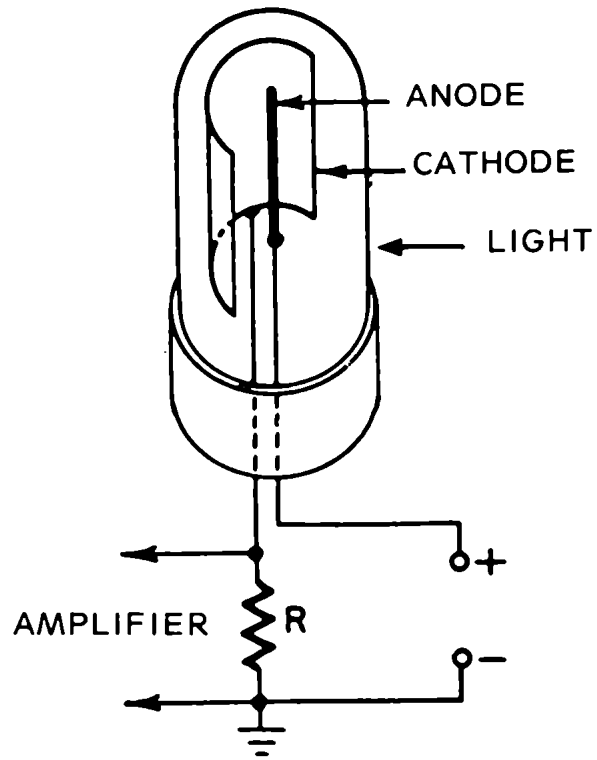


Fig. 2-38. Drawing of a photoemissive cell.

PHOTOELECTRIC TRANSDUCERS

There are a number of devices in which the electrical characteristics change upon illumination or irradiation. Such devices are serviceable as photoelectric transducers. We examine photoemissive cells belonging to electron tube technology and solid-state devices.

Vacuum & Gas-Filled Tubes

Photoemissive cells are cold-cathode devices. Electrons are emitted when light (photons) strikes the cathode. Such a cell comprises an evacuated glass bulb containing a large generally curved metal plate which acts as a cathode, and a stiff wire or a wire loop, the anode, placed between cathode and light source (Fig. 2-38). To intercept as little radiation as possible, the anode is loop shaped. By a suitable DC voltage, the anode is made positive in respect to the cathode, and load resistor R is inserted in the circuit. Upon illumination, the cathode emits electrons that are attracted by the anode. The current thus formed develops a voltage drop across R, which is fed to the input of a suitable amplifier.

The cathode base metal is covered by a strongly photoemissive material (i.e., a material which emits as many electrons as possible for a given number of impinging photons), generally a mixture of alkali metals that determine the spectral response of the cell. The cesium cathode, overlaid by a silver-doped cesium oxide film, is most sensitive to red-infrared radiation, in conformity with standard spectral sensitivity characteristic S1. Using an appropriate optical filter placed before the cell to filter out visible light, a device responding only to (invisible) infrared radiation is obtained. The cesium-antimonide cathode features a peak sensitivity to ultraviolet radiation (standard characteristic S4). There are still other cathodes approaching the spectral sensitivity of the human eye.

Even without any radiation, there still is a residual current, called **dark current**, due to the thermionic emission of electrons. This current increases with ambient temperature; therefore, it can be lowered by cooling the cell, and this is done in some scientific and military applications. In industrial control functions, however, the light level is sufficiently high to render negligible any dark-current effects.

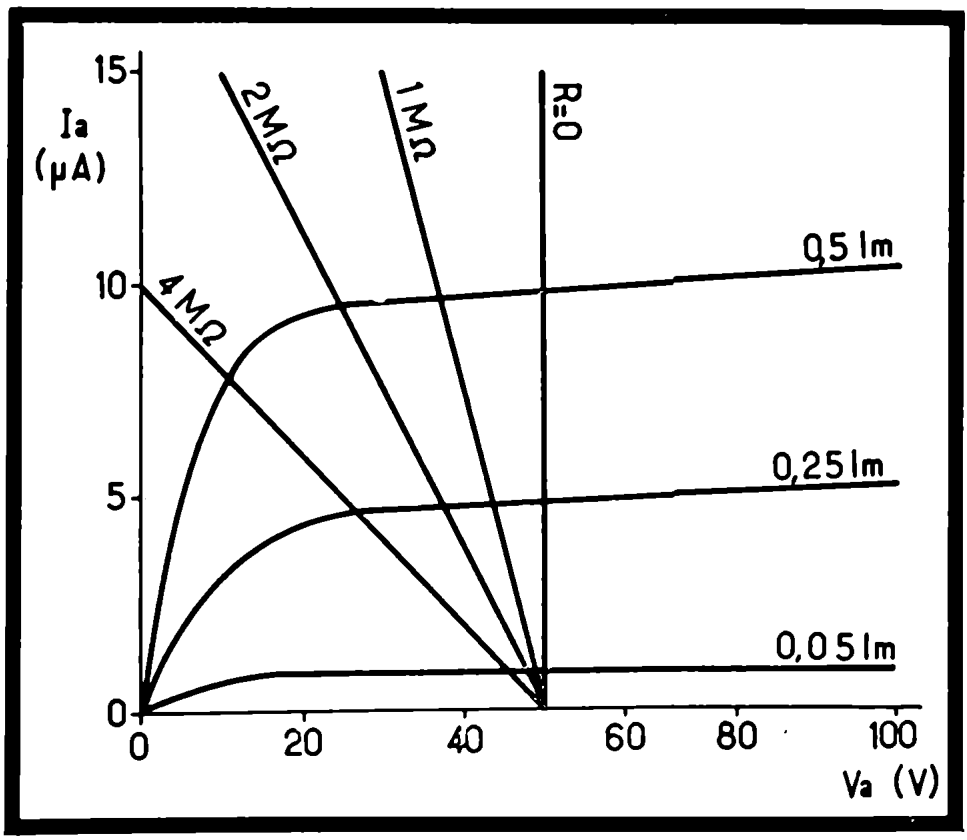


Fig. 2-39. Graph showing the illumination-dependent characteristics of a vacuum photoemissive tube.

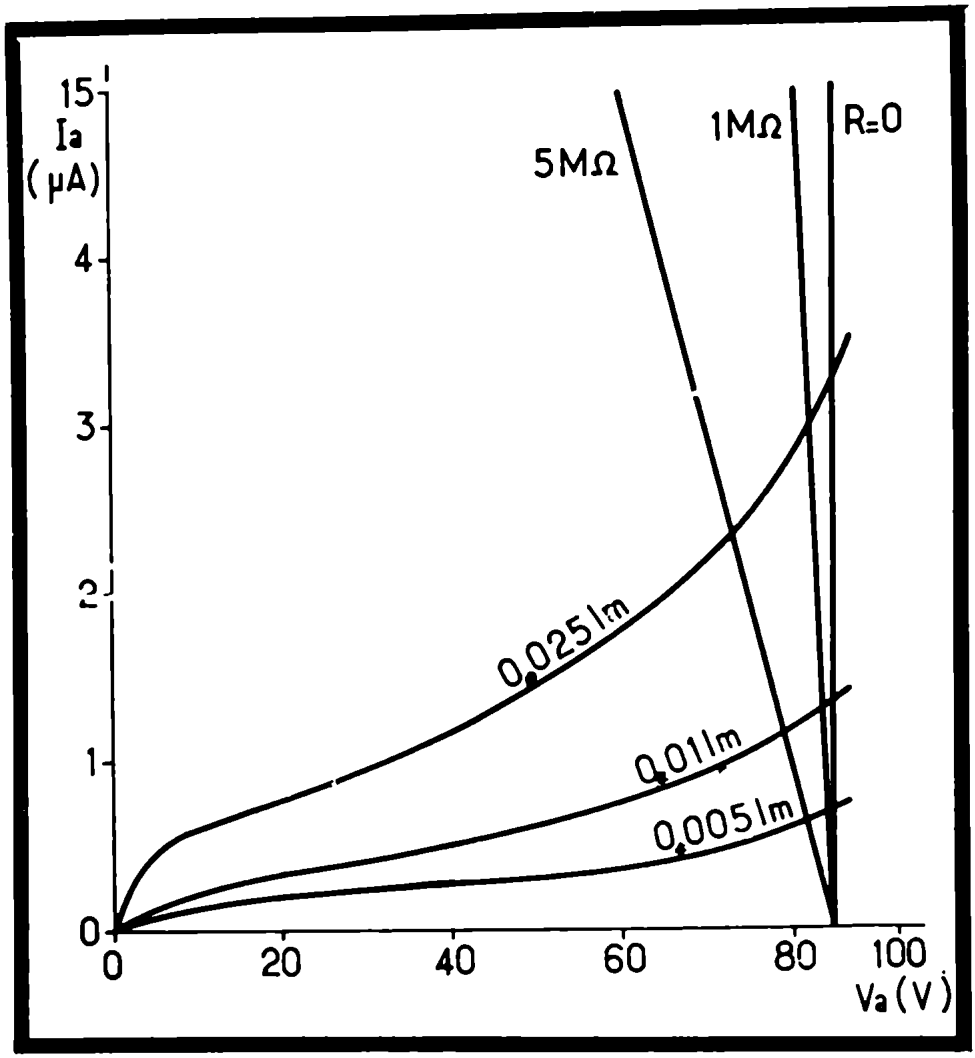


Fig. 2-40. Characteristic curves for a gas-filled photocell.

The voltage-current (Ia over Va) characteristic of a photoemissive cell is shown in Fig. 2-39 for brightness levels of 0.05, 0.25 and 0.5 lumen. For an anode voltage exceeding 20 to 60v, saturation sets in, and Ia varies very little with Va. Furthermore, Ia is small, less than 0.1 ma, so that a high-impedance input amplifier is necessary. The sensitivity of this transducer is about 20 to 50 microamperes per lumen.

In the evacuated envelope, some rare gas such as argon, xenon, krypton or helium is introduced. The result is a gas cell behaving in a very different way (Fig. 2-40). Electrons leaving the cathode meet gas atoms. The gas atoms split into positive ions proceeding toward the cathode. Other electrons join the originally emitted electrons to impinge upon the anode. The result is a "gas amplification" factor of about 5 to 10, bringing the overall sensitivity to about 100 to 150 microampere per lumen.

Considering the characteristics, a slight knee corresponding to the saturation segment is noted, but then Ia increases more and more; exceeding an anode voltage of about 90v, the phototube becomes a glow tube and is definitely destroyed. Gas tubes, therefore, must be protected against a peak anode voltage in excess to that specified.

Vacuum-type photocells feature an anode current proportional to brightness, and their operation is reproducible. Therefore, they are well suited for photometric (analogic) measurements. This proportionality does not hold for gas-filled cells in which the sensitivity decreases somewhat during operation. For these reasons, they are definitively unsuited for light level measurement purposes. On the other hand, they are quite serviceable in on-off-type controls. In modern equipment, at least, the gas-filled cells are superseded by semiconductor devices featuring much greater ruggedness and compactness. There is one type of photoemissive tube that cannot readily be replaced by a solid-state device: the **photomultiplier**, a sensing device for very low radiation levels for astronomic, nuclear and other scientific applications.

The elements in a typical photomultiplier are shown in the drawing in Fig. 2-41. The photocathode is made up of a translucent film of S4-type photoemissive material (cesium antimonide) which coats the plane faceplate of the glass envelope. Electrons emitted by the cathode are attracted by the first anode called **dynode** 1 (D1), arranged so that every incident electron sets free several secondary electrons to proceed together toward dynode D2. D2 is set at a higher potential to attract electrons originated by D1. There, again, electron multiplication takes place, and the same pattern repeats several times more because there are 11 dynodes. A global amplification of 106 and up thus is obtained, and the device sensitivity reaches 125 amperes per lumen! (This, of course, does not mean that you get an output of 125 amps if you put in 1 lumen, because the tube will be saturated long before, but for 1 microlumen, you may get 125 microampere.) The dark current remains lower than 5×10^{-8} ampere; therefore, the device can receive light pulses a hundred times smaller than a photocell followed by an amplifier.

Because of its higher signal-to-noise ratio (the signal stands out better within the noise) and of its lower time constant, a single photomultiplier is to be preferred to a photoelectric device followed by an amplifier of similar gain. On the other hand, a dynode typically requires about 150v, and because of the number of cascaded dynodes, a total anode voltage of about 1,800v, depending on the tube type, is

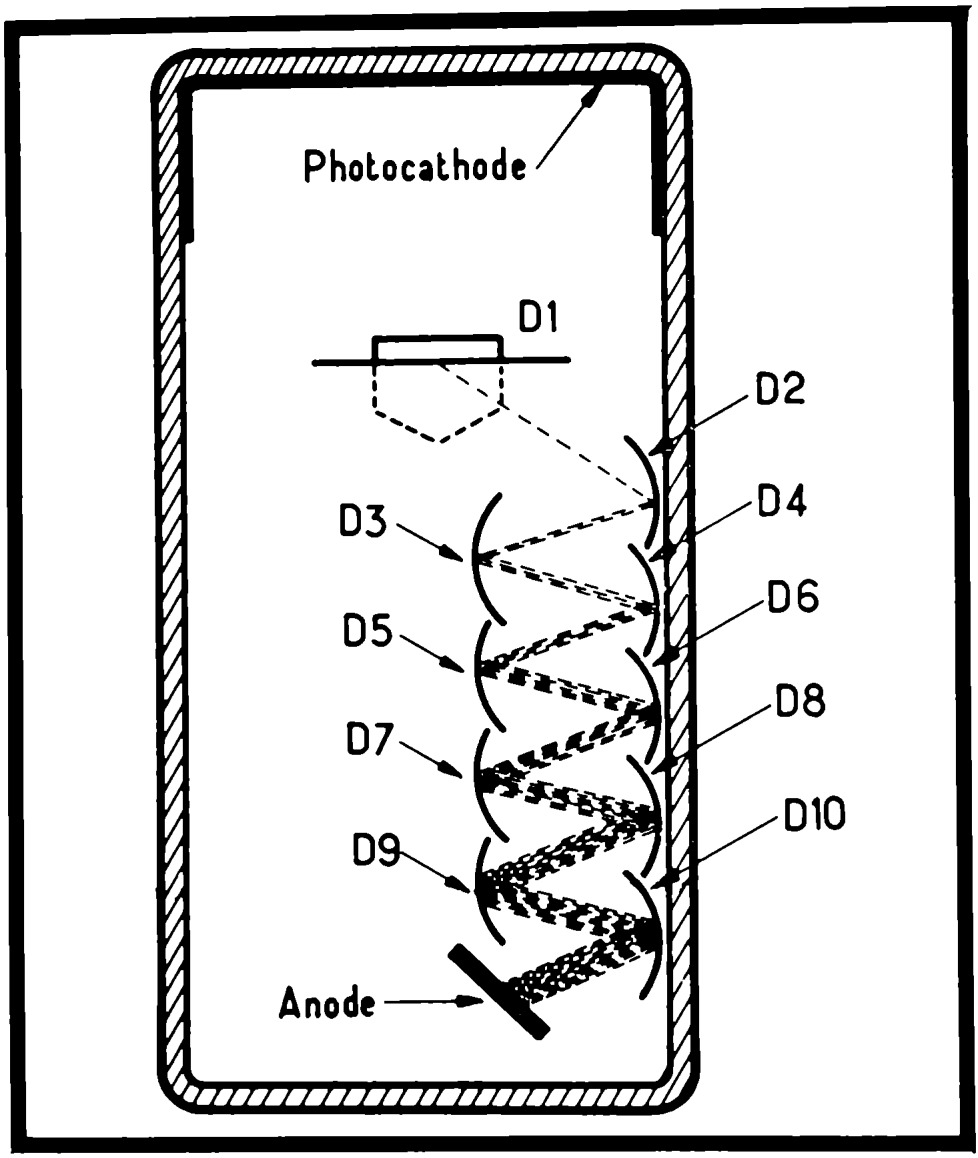


Fig. 2-41. Drawing showing the electrode arrangement in a photomultiplier tube.

necessary. Furthermore, this voltage must be stabilized since the highest electron multiplication is obtained for a definite set of dynode operating voltages.

Solid-State Devices

Photoelectric semiconductor devices fall into photovoltaic cells, photoconductors and junction devices. Among all photoelectric devices, the **photovoltaic cell** has the unique property of developing an electromotive force when stimulated by light. The older batteryless photographic exposure meters contain such a cell made of the selenium which supplies a built-in galvanometer with a current that increases

with the light level. The solar generators powering satellites also contain photovoltaic cells.

The energy conversion efficiency of conventional selenium cells is rather low (about 1 to 2 percent), but there are silicon cells featuring an efficiency of 10 percent and more. Because of the delicate processing, silicon cells are expensive in comparison with selenium cells. However, as a means of providing more power to space flights, the added expense is more than worthwhile. And even on earth, low-power remote equipment can thus be powered economically (considering other solutions).

A section of a selenium cell is shown on Fig. 2-42. Selenium is deposited on a metal base and covered by a film of a metal such as gold, sufficiently thin to let light shine through it on the selenium. A metallic frame contacts this translucent electrode. The sensitive side is generally covered by a glass window or a transparent resin for protection.

A cross-section of a silicon cell appears in Fig. 2-43. The surface to be made light-sensitive is composed of a slice of n-type silicon that is diffusion-treated to form an extremely thin p-layer, so thin that light can pass through it and reach the p-n junction. An electromotive force develops upon illumination of this junction; in bright sunlight, it generates about 0.5v. The p-layer is positive, of course. Since the available current is limited, a number of such cells are series and parallel connected to yield a useful output.

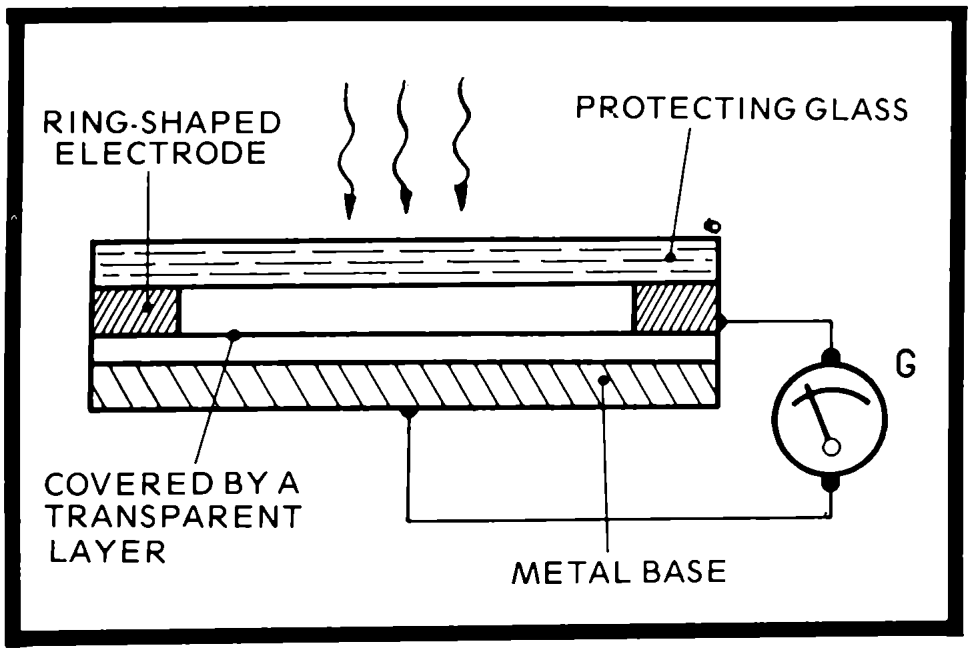


Fig. 2-42. Cutaway view of a selenium photovoltaic cell.

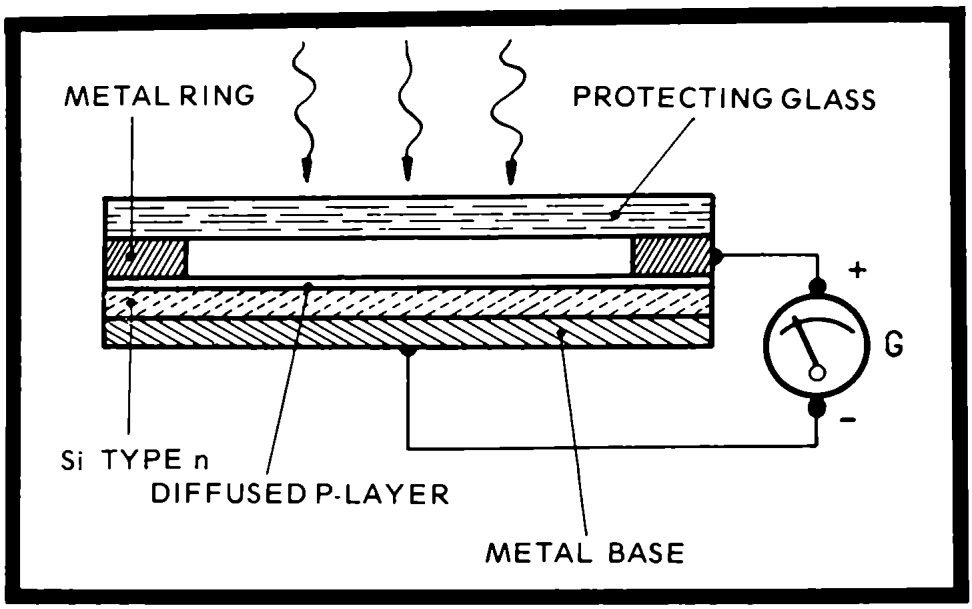


Fig. 2-43. Drawing of a silicon photovoltaic cell or "solar battery."

Photoconductors are made of amorphous semiconductor materials such as cadmium sulfide (CdS), cadmium selenide, selenium and lead sulfide (infrared sensitive). These materials feature a very high dark resistance that drops to a fairly low value upon illumination; ratios such as 10,000:1 are common. Producing no energy, photoconductors must be supplied with power to work. Photographers know that the CdS-exposure meter of their camera, while being a lot more sensitive than the former selenium cell, depends upon a battery to work.

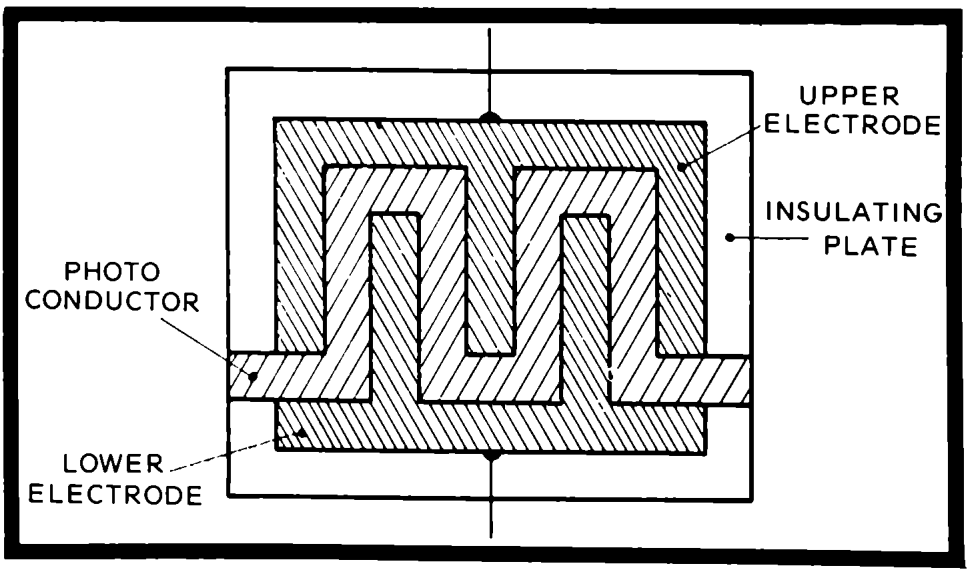


Fig. 2-44. A photoconductor is a light-sensitive resistor.

To prepare a photoconductive cell, two interdigital metallic electrodes are sputtered upon a dielectric substratum, and the meandering path between them is filled with the photoconducting material (Fig. 2-44). By dimensioning the active cell area, the device can be designed for power ratings of, say, 5 mw to 2w. The latter cell is able to control 2w of power, enough for a standard relay, thus allowing for the simplest possible control circuits. Such large-area cells, however, are often inconvenient. For reduced bulk requirements, pinhead-size bead types are chosen; the low rating is made up by a transistor or thyristor stage. While photoconductors are often used in low-voltage circuits, there is no junction to break down, and particularly the power types resist applied voltages of 300 to 500 (provided the rated power is not exceeded).

Photoconductors do not respond instantly to an exciting light. A rise time of about 5 ms for high-level and as much as half a second for low-level illumination can be anticipated, and a similar delay occurs as the light extinguishes. This limits the use of these sensors to somewhat slow phenomena. For repetitive events, 1000 Hz is about the frequency limit for fair light level.

Photoelectric **junction devices** are semiconductor diodes and transistors with a housing containing an opening that holds a tiny convex lens to concentrate incoming light rays on the junction. In a photodiode, the junction is reverse-biased, so it passes only negligible current. Illuminating the junction increases its reverse current which becomes the working current of the device. (To prevent this photoelectric effect in ordinary diodes and transistors, the semiconductor is housed in an opaque case..)

There are germanium and silicon photodiodes; the latter are less temperature-sensitive and their dark current is lower. These are the smallest photoelectric transducers available, some types are match-size, with the little lens on one end. A number of these diodes may be positioned side by side in optical readers. They operate at frequencies much higher than other light-sensitive devices: up to 100 kHz for germanium and 1 MHz for silicon. The working current is less than 1 ma, and an amplifying stage is normally required.

The **phototransistor** is a sort of photodiode with a built-in amplifier. It differs from a conventional transistor by a glass window in the top of the case to let light strike directly on the emitter-base junction. Just like injecting current in the base, illuminating this junction results in an amplified collector current sufficient for relay operation. Thus a single

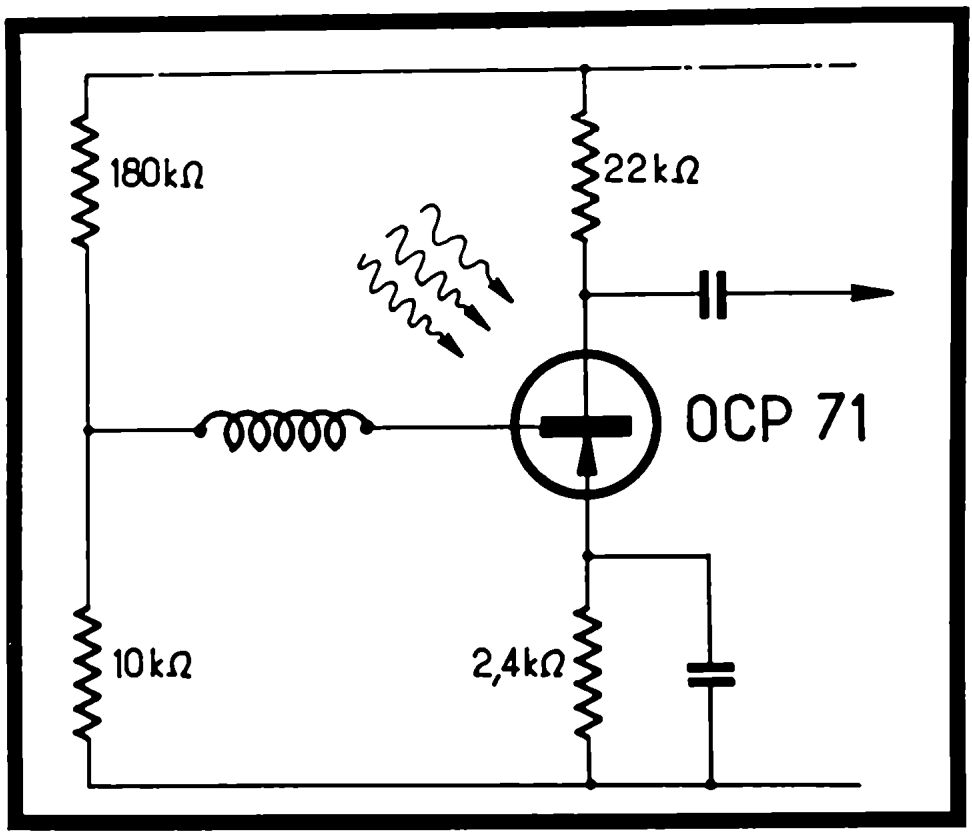


Fig. 2-45. The emitter-base junction in a phototransistor is allowed to response to light.

phototransistor can supersede a photodiode followed by a transistor.

A phototransistor connects just like a common transistor (Fig. 2-45). Being light-controlled, the device lacks an electric signal input (though such a signal can be applied to the base). But since a varying temperature has the same effect on collector current as a variation of incident light, the base may be tied to a voltage divider for thermal stabilization purposes. However, if the device is operated as an on-off control, there may be no point in stabilizing the dark current, since the two states are sufficiently differentiated and the base is left free. A 2-wire device like the photodiode is the outcome.

Because of the built-in amplification, the phototransistor is rather more light-sensitive than the photodiode: about 300 ma per foot-candle, against 2.5 to 20 for the latter. Being made of the same materials, both devices are responsive to the same visible light and near-infrared parts of the spectrum.

The light-activation principle obviously also works with other semiconductor devices such as silicon controlled rectifiers (SCR), silicon controlled switches, thyristors, etc. As these are power-control devices (examined later), a tiny

component is able to switch on and to cut out loads of up to 100w by means of a simple light beam.

OPTO-ELECTRONICS

This term designates a domain comprising electronic circuits having at least one coupling or feedback by optical means such as light rays. This calls for provision of a photoelectric transducer for converting information-bearing light variations into electrical signals. According to this definition, the long-known photoelectric controls are opto-electronic devices. However, giving a new name to a well-known technical domain always stimulates development of new devices and applications.

An optical relay is composed of a transmitter and a receiver of optical radiation and of the medium transmitting the radiated energy (Fig. 2-46). While the human eye senses only the small portion called "visible light" of the radiation spectrum (wavelengths between 0.4 and 0.7 micrometer with an utmost sensitivity for greenish yellow at 0.55 micrometer), the various transmitters and receivers available more or less exceed these limits on the high side (wavelengthwise) towards the (invisible) infrared (more than 0.7 micrometer), or even on the low side toward ultraviolet (less than 0.4 micrometer) and, more remotely, X-rays. For highest transmission efficiency, transmitter, transducer and transmission path medium are obviously to be worked at their preferred wavelength.

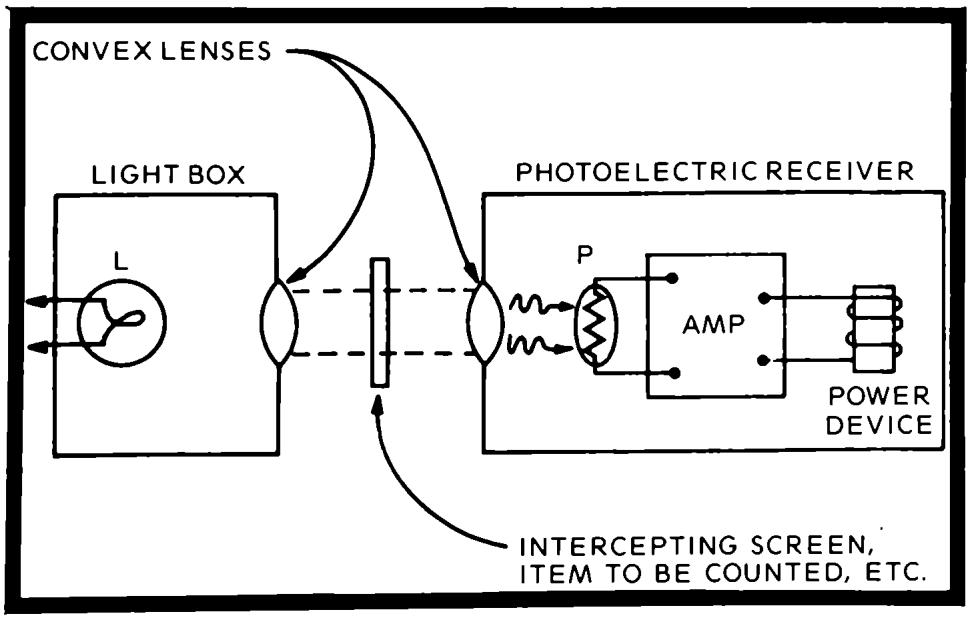


Fig. 2-46. A photoelectric relay is composed of a light box and a photoelectric receiver.

The most common electric light source is the common **incandescent lamp** which emits a yellow-red to infrared radiation. Its service life can be considerably lengthened by a slight underrating, so that reliability may be acceptable. Because of thermal inertia, the heated wire, however, is unable to follow any but the slowest voltage variations, limiting its use to almost steady light sources. The **gas tube** (mostly neon or argon filled) starts and extinguishes almost instantly and can be operated at audio frequencies; it emits a yellow-red light according to the gas mixture. Its ionization voltage generally exceeds 55 volts, but it can be reduced by suitable biasing to make transistor control workable.

Solid-state **electroluminescent diodes** using semiconductor materials such as gallium arsenide-phosphide (Monsato) and gallium phosphide are the latest red color radiation emitters the opto-electronician is presented with. This diode works inertialess and heatless, its life is almost unlimited, and its ruggedness, of course, is much higher than that of other lamps. And it is compact, too; a recent Monsato device looks like a glass droplet with two connecting wires. The emitted-light wavelength of Model MV 50 is centered on 0.65 micrometer; it needs but a few volts to work and takes 20 ma to emit a very bright light (750 foot-lamberts). By its circuitry, and even by its production technology, this device is completely compatible with integrated circuits.

Photoelectric transducers have been described. So the third and last element of an optical communication system is the **transmission medium** between the source and the receiver. It is well known that light waves travel best in air, and that opaque screens, even very thin, as well as rain, snow and fog more or less absorb the radiated energy and thus impede the communication. The ease of interception of a light beam, however, can be made a virtue by designing photoelectric controls operated contactlessly by intercepting or restoring the sensor illumination for counting and safety purposes.

Secure, weather-independent communications in air are possible only if the medium is protected by enclosing the transmission path in some sort of pipe, thus keeping out obstructing particles. Although such a proposition is not very practical, it directly leads to the important concept of **light ducts** and **fiber optics**. Take a transparent glass or plastic rod and shine a light into one end; the light will travel within the rod and can leave only at the other end; because of the differing refraction indexes of the glass-air interface, total reflection takes place, and this interface just acts as if there were a pipe with mirrored inner wands. And you can gently

bend and even knot the rod, and the light will keep on leaving on the sole output face. This feature allows for "piping" light from one point to another, passing around obstacles. Thanks to it, the source and detector do not necessarily face each other.

If there is more than one light signal to be transmitted, a number of transparent rods can be assembled into a row. In fact, the rods may be fine fibers, and there comes "fiber optics," as illustrated by Fig. 2-47. To prevent transmission of light from one fiber to the next one, each fiber is covered by a glass or plastic sheath of greater refractive index to provide total reflection at the interface. It is this sheath that materializes the pipe wands containing the "piped" light. Since the individual pipes can be very fine, a cluster composed of a great number of fibers can be grouped together. Both ends of the string thus obtained are then trued up and finely polished. If one end of this optical device is placed on a paper, the print it recovers appears on the free end and looks as if it were in the output plane.

PHOTOELECTRIC CONTROLS

As outlined above, there are a good many serviceable photoelectric transducers, and the designer has to look for utmost reliability and only thereafter for economy. Gas-filled phototubes are obsolete, and vacuum-types are limited in use to measurement purposes only and for the detection of very weak signals. Barrier-layer cells are seldom used in industrial control. So the choice is limited to photoconductors and photosensitive junction devices of all kinds.

Photoconductors lead to the simplest possible circuit, for they are resistors without polarity and thus can be powered by raw AC. However, this may be inconvenient for powering a relay, generally designed for DC operation; when AC is ap-

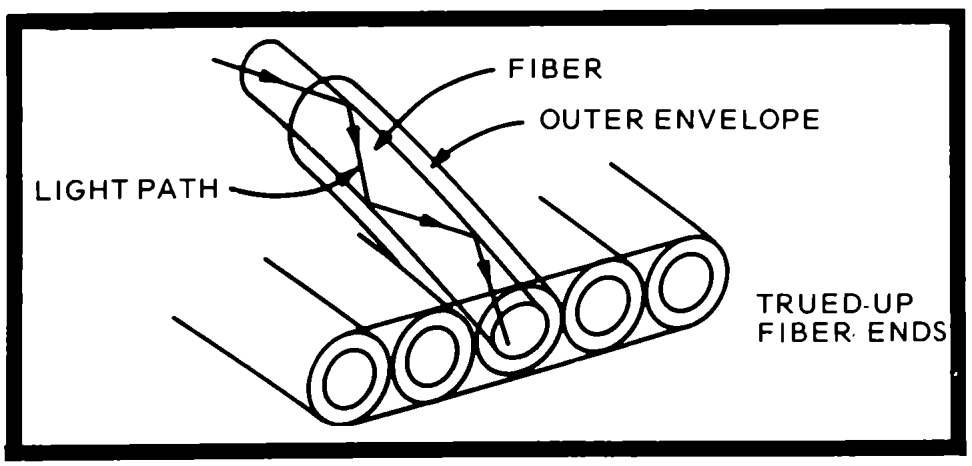


Fig. 2-47. An optic fiber behaves as a light duct or guide.

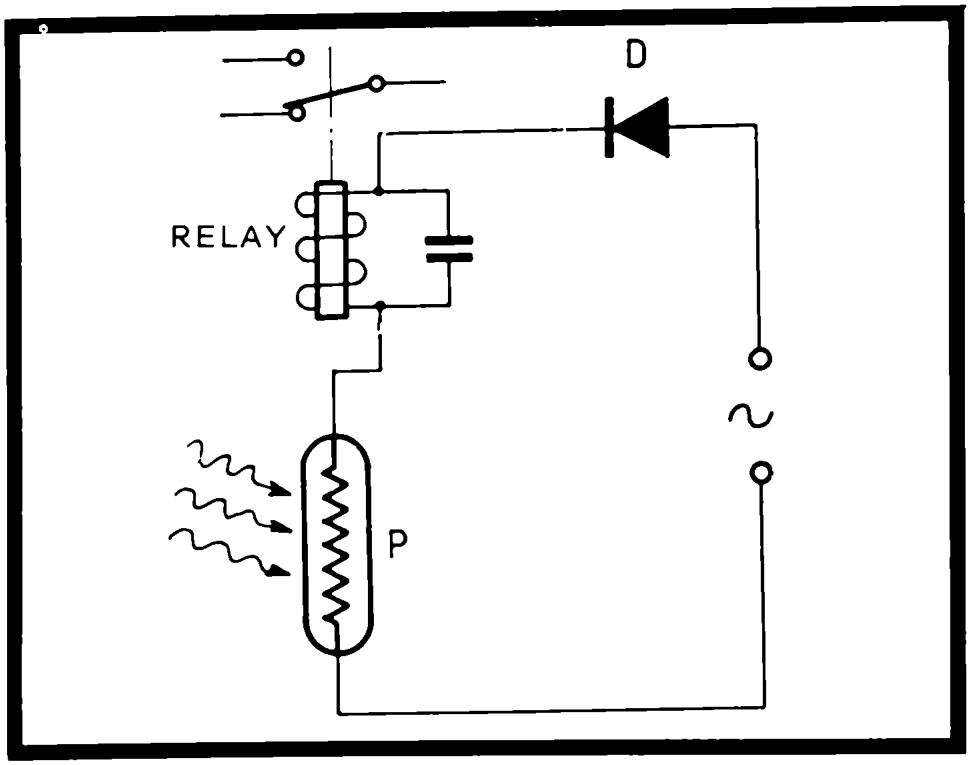


Fig. 2-48. A large-area photoconductor can directly control a sensitive relay.

plied, such a relay will get warm and its armature will vibrate, making for noise and possibly bad contacts.

The circuit in Fig. 2-48 shows a diode rectifying the AC supply; capacitor C is to reduce relay armature vibration due to half-wave operation. In the dark, resistance of photoconductor P is high, and the relay is not energized. Upon

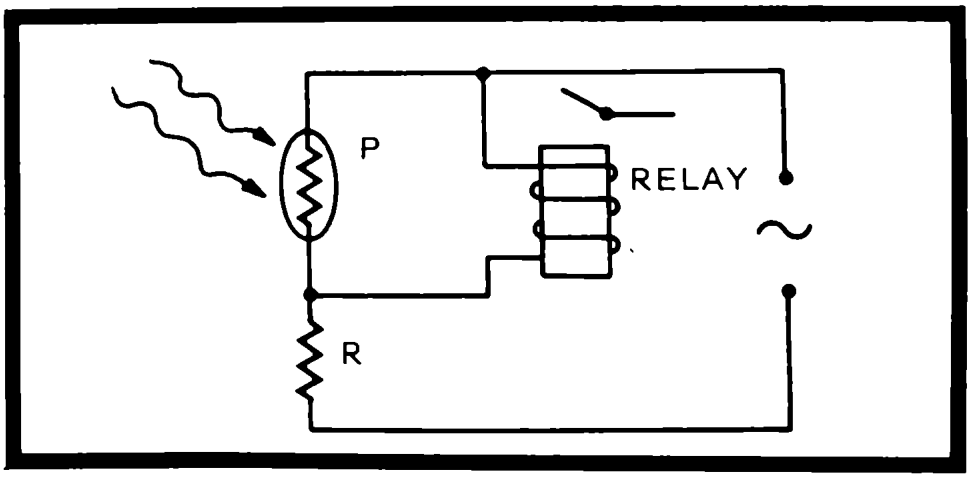


Fig. 2-49. In this circuit, when the photoconductor shunts the relay, illumination energizes P and de-energizes the relay.

illumination the resistance of P is lowered significantly (to, say, one one-thousandth of its original value), and the increased current is sufficient to operate the relay. This is a series circuit and the relay operates upon illumination. A parallel circuit in which the relay is energized in darkness or on illumination interruption is shown in Fig. 2-49. An AC relay is used here, making a rectifier and capacitor unnecessary. On the other hand, series resistor R forms a voltage divider with P. In the dark, the resistance of P is high, and the relay, virtually unshunted, is energized. Impinging light drastically lowers the resistance of P, bypassing the relay whose armature falls open.

Such power photoconductors can directly control relays or even small motors; but because of the large area necessary, their application is somewhat restricted. The circuit in Fig. 2-50 uses a small-size photoconductor fed by a reduced voltage (12v) followed by a transistor amplifier, or rather a Schmitt trigger in which the flop-over makes for clean relay operation. All the circuit couplings are resistive and resistor R1 is the operating threshold adjustment.

If we replace T1 in this circuit by a type housed in a case with a small window to illuminate the emitter-base junction (that is, a phototransistor), we obtain the circuit in Fig. 2-51,

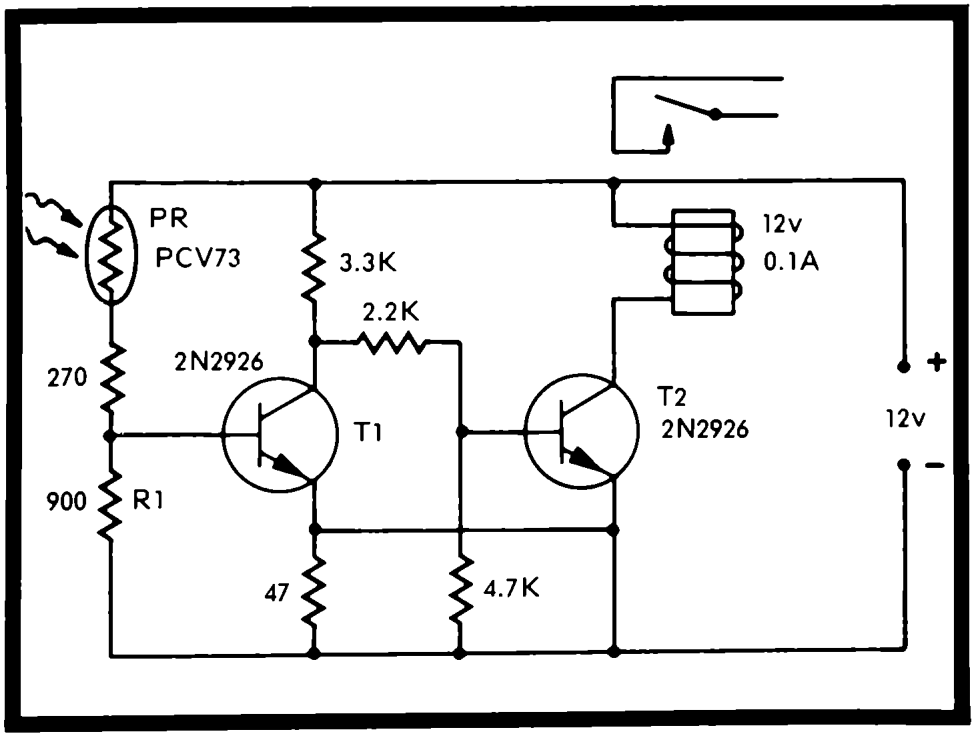


Fig. 2-50. Schematic of a photoconductor-controlled Schmitt trigger.

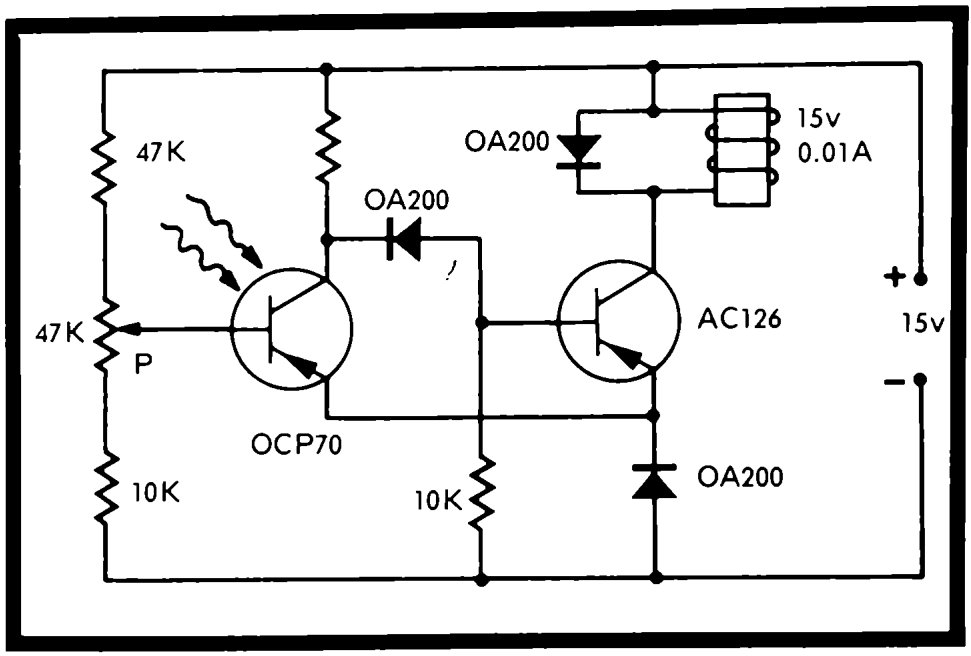


Fig. 2-51. Schmitt trigger circuit incorporating a phototransistor.

where a flop-over again is activated by illuminating a photoelectric device. The relay now is energized upon vanishing illumination. Potentiometer P allows for temperature and ambient light compensation.

The major feature of a photodiode is its compactness, essential to a number of applications, but its photocurrent is too low to be used directly. A very simple circuit comprising a photodiode followed by a common-collector transistor (Texas

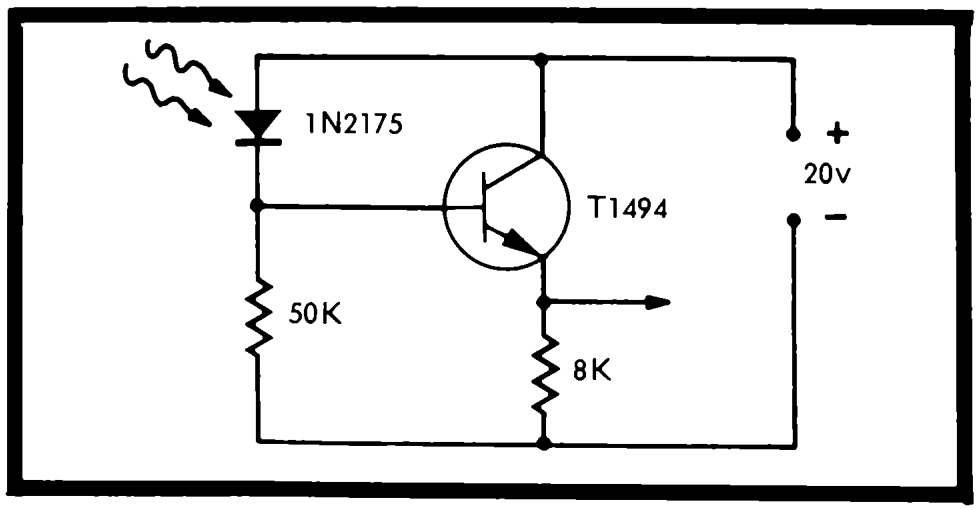


Fig. 2-52. Circuit showing a photodiode coupled to a common-collector transistor amplifier.

Instruments) featuring a low-impedance output is shown in Fig. 2-52. A single phototransistor should yield the same result without any coupling problems. The germanium photodiode in Fig. 2-53 is followed by a Darlington-connected preamplifier. The thermal dark current variation of the photodiode here is almost entirely overcome by the thermal rest current of the first transistor, thus the circuit is capable of withstanding large temperature variations without noticeable drift. The high input impedance perfectly matches the high diode impedance and the low output impedance makes for easy utilization. Peak output is about 5v.

Phototransistors are often used (especially in on-off, or digital circuits) as photodiodes, leaving the base free (Fig. 2-54). As the thermal compensation is impossible under these circumstances, resistor R and the diode form a voltage divider in which R is made adjustable. The relay control amplifier is unidirectionally coupled by diodes to pass only one-polarity variations.

Thyristors and similar devices, discussed in the following chapter, can be established for light-beam control, at least the lower power types. For instance, the General Electric LASCR (Light Activated Silicon Controlled Rectifier) is nothing other than a thyristor where the tophat-housing is fitted with a tiny lens to illuminate the photosensitive junction area of the silicon chip. There are also light-activated 4-layer devices. As these "solid-state switches" control considerable power by on-off action, they can directly energize relays, electromagnets or motors, leading to circuit simplification.

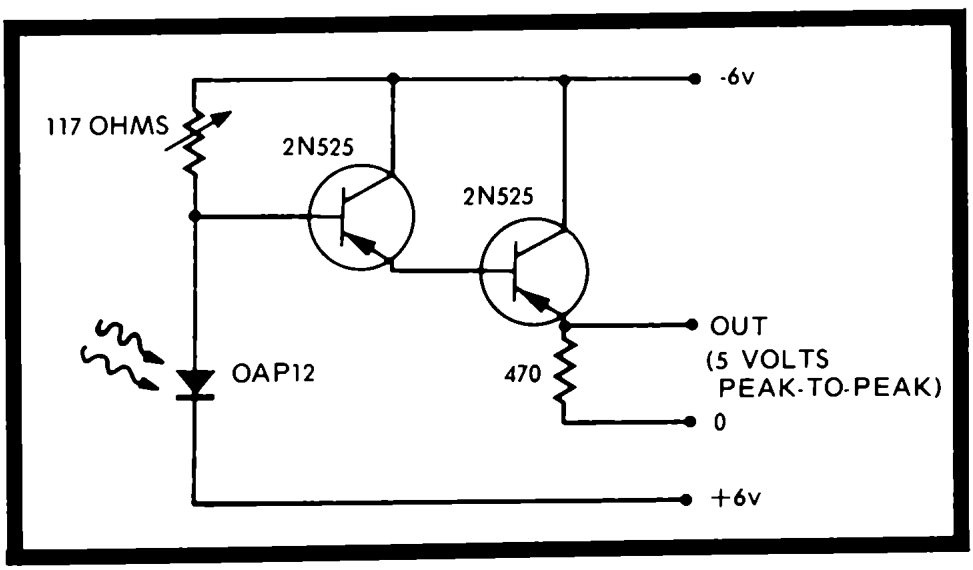


Fig. 2-53. Schematic of a photodiode connected to a Darlington pair.

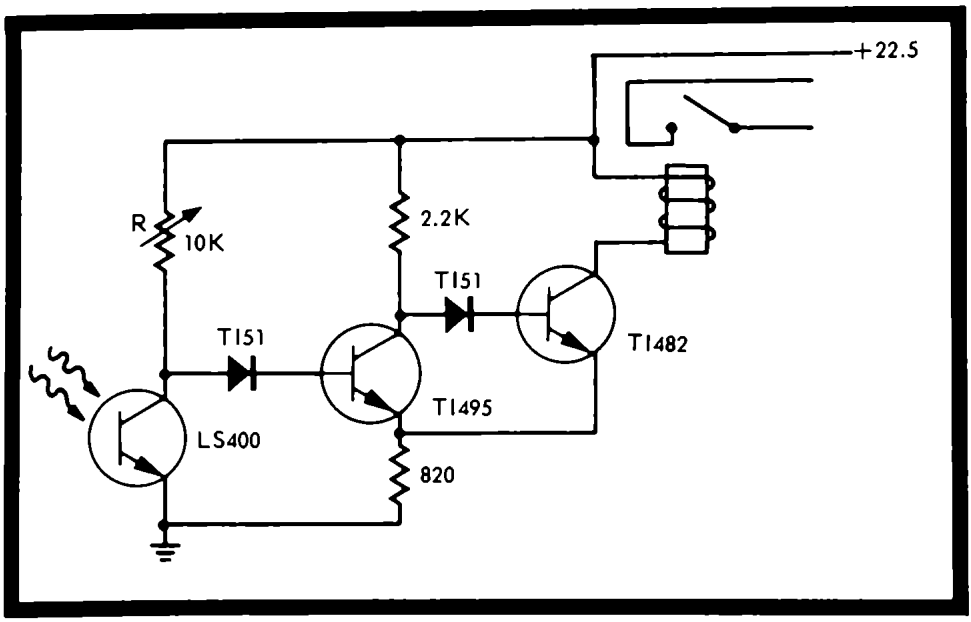


Fig. 2-54. Phototransistor used as a low-impedance photodiode.

Since the photoelectric relay in Fig. 2-55 is fed by AC, the photothyristor can pass current only on half-cycles of one polarity; it becomes nonconductive during the following half-cycle and can be made to conduct during the next if the light signal is still present. A standard DC relay is used, paralleled by capacitor C to avoid vibration. Resistor R, series-connected with C, limits the transient charging current of C to a safe value and thus protects the thyristor.

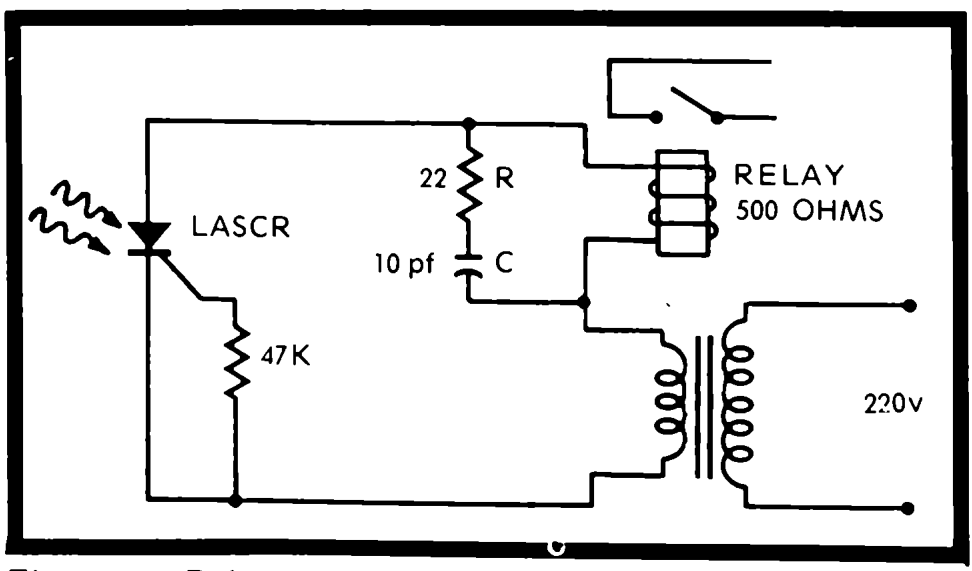


Fig. 2-55. Relay circuit controlled by a light-activated thyristor (LASCR).

Present photothyristors can stand only comparatively low supply voltages, less than about 50 RMS volts. For a transformerless 220v main power line operated photorelay, the circuit in Fig. 2-56 uses a thyristor (Ty) of adequate rated voltage to energize an AC relay laid-out for 220v RMS. Ty is controlled by a 4-layer photodiode (Telefunken) connected between a bias voltage divider and the diode gate. Unlike the silicon controlled rectifier (SCR) which has but one gate, the 4-layer device sports an anode gate Ga and a cathode gate Gc. While Ga often remains unconnected, Gc is always tied to the cathode through a resistor to avoid accidental firing of the device by main power line transients or by inverse current in the case of a temperature rise.

In the following pages, specific photoelectric device applications are discussed. Therefore, we will illustrate two basic principles often used in the design of photoelectric controls. The first one is the **differential control** represented in Fig. 2-57. There are two light-sensitive transducers PC1 and PC2 coupled to a differential amplifier. A unique light source, L, illuminates both cells by optical means such as mirrors, prisms, lenses and-or optical fibers. Thus a brightness variation of L will act similarly on both channels and cancel at the output.

Furthermore, this combination forms an opto-electronic bridge in which the variable arms are transmission mediums X and Y located between the common light source and each photocell. It is thus possible to measure the opacity differences of two translucent screens or of two liquid solutions

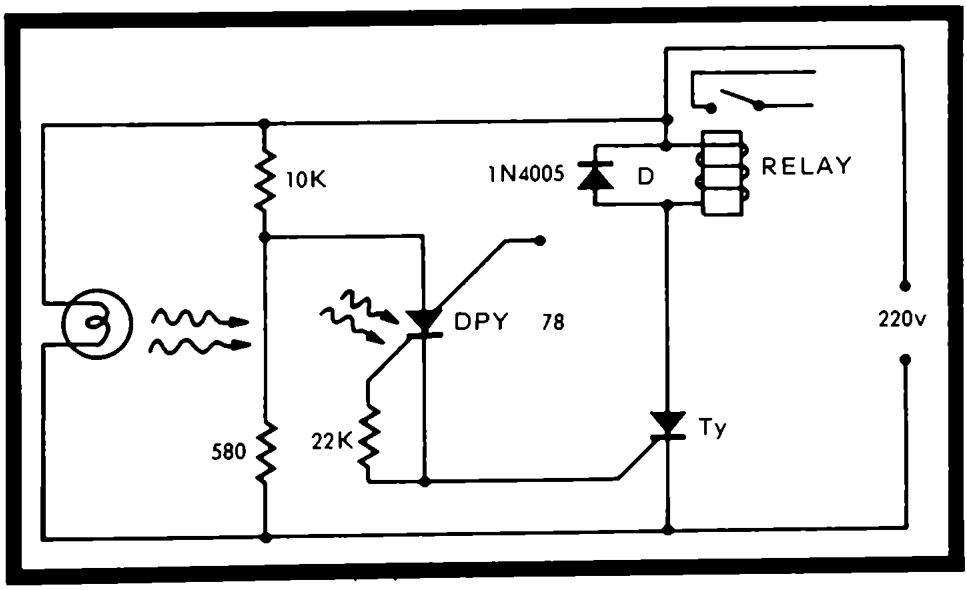


Fig. 2-56. The high-voltage thyristor in this circuit is controlled by a 4-layer photodiode.

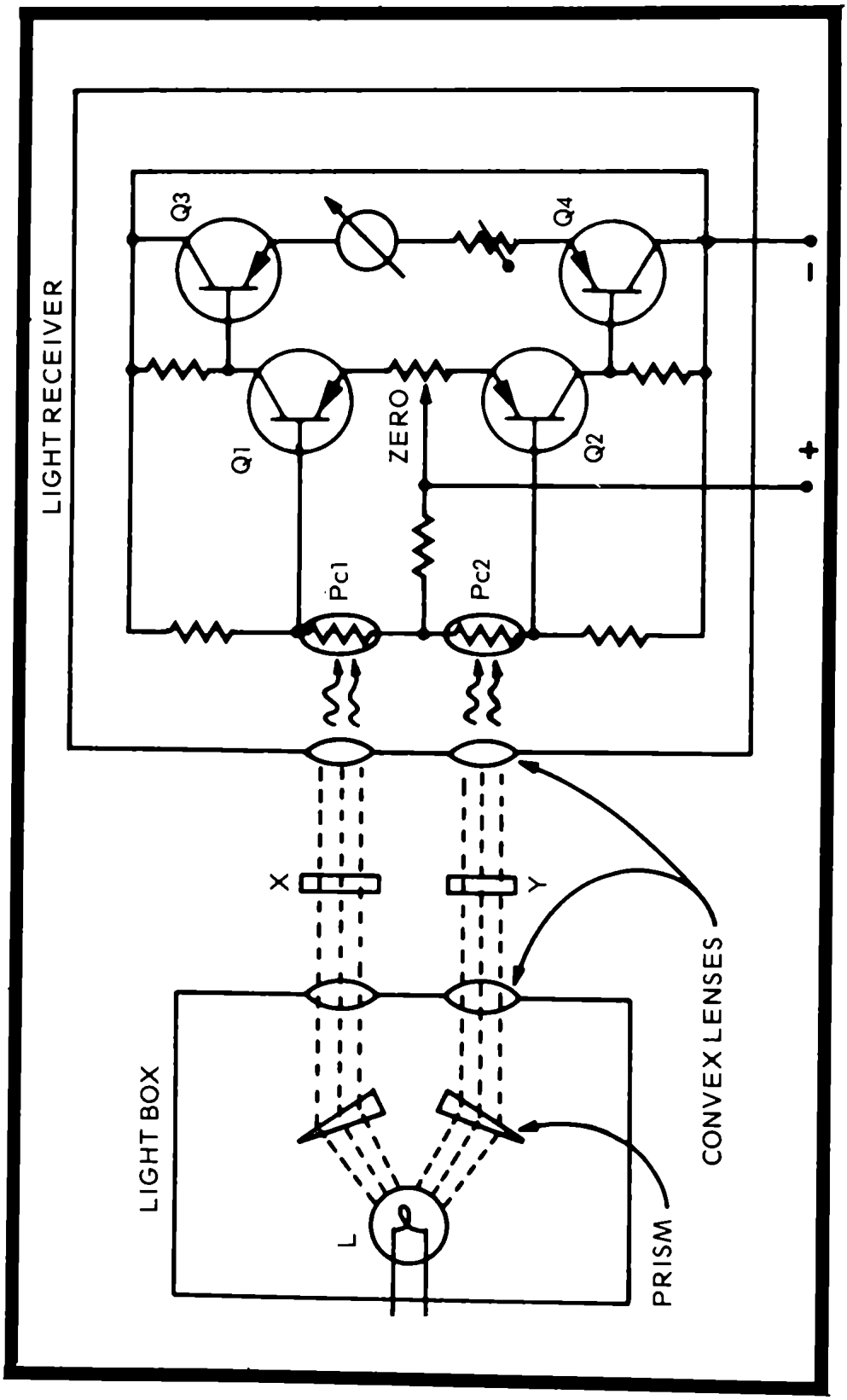


Fig. 2-57. Schematic of a differential light-sensitive control.

contained by transparent tanks placed at X and Y, or even to detect the presence of smoke intercepting the light transmitted by only one channel; the other is protected by suitable screening against outer interference.

Another important problem is to make the useful light "signal" stand out amidst the ambient light that tends to swamp it. Be it daylight or artificial, the ambient light is of a continuous nature; to be easily distinguished, the light signal has to be alternating and preferably pulsed. This is obtained mechanically by a **light chopper**, an opaque wheel perforated by a number of equally-spaced holes and rotating at a definite speed. This wheel is positioned so that the light beam issued by the source has to pass by one of the holes to reach the photocell (Fig. 2-58). Suppose this wheel rotates at a speed of 1,500 rpm, or 25 rps, and that there are 20 holes; there will be 500 light pulses per second, or a repetition frequency of 500 Hz. The AC amplifier fed by the photocell signal can be laid-out (preferably by RC networks) to selectively transmit that frequency, so that only the chopped signal will produce an output. Of course, the wheel speed has to be unvarying to get the signal through, possibly by using a synchronous motor.

But the wheel system is subject to wear and requires some maintenance; and then, static devices nowadays are preferred wherever possible. A gas tube or flash tube fed by an audio frequency voltage of sufficient amplitude may form such a static pulsed-light source. However, the electroluminescent gallium arsenide or arsenide-phosphide diode mentioned above should qualify as the best possible source of this kind for this type of application. Working with the same voltages and currents as transistors, it can easily be excited by a low-voltage oscillator, and even is compatible with integrated circuits. Furthermore, it generates, with a high yield, a red and infrared radiation in the high-sensitivity region of most solid-state transducers.

OPTO-ELECTRONIC LINK OR MODULATOR

Assembling an emitter and a transducer for the same type of radiation into a compact 4-wire device results in a new component usable as an opto-electronic link in a closed-loop system and also as an on-off modulator or chopper. These devices feature a nearly perfect galvanic insulation between input and output circuits, justifying the "Photon-Coupled Isolator" trademark of one manufacturer.

The **Raysistor** (Raytheon) was an early opto-electronic device composed of a tiny lamp facing a photoconductor contained in the same small cylindrical housing (A, Fig. 2-59).

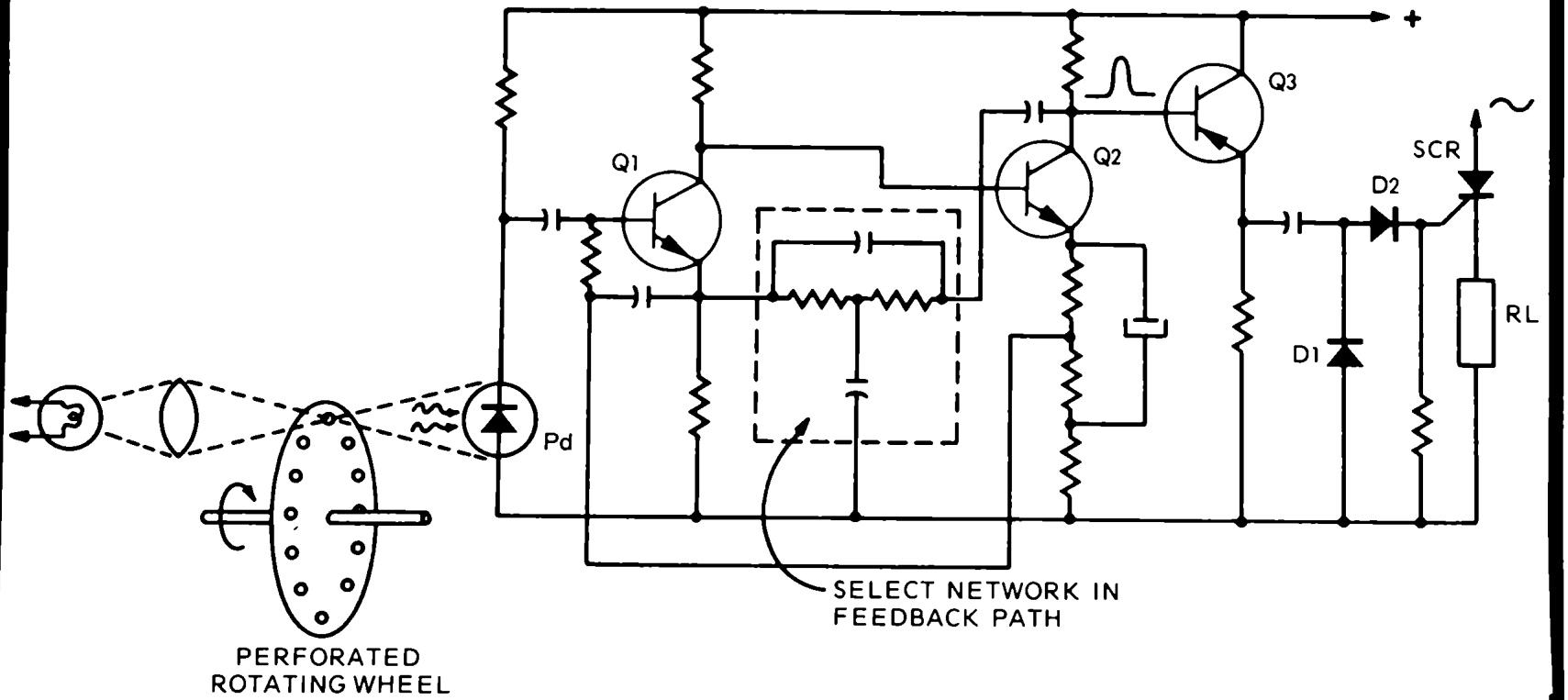


Fig. 2-58. A mechanically chopped "signal" light is passed by the tuned amplifier, while the steady ambient light isn't.

In another Raysistor type, the incandescent-wire lamp is replaced by a glow lamp (Fig. 2-59B), thereby increasing the working frequency limit from the very low to medium-high audio frequencies. More recent opto-electronic modulators such as the TIXL 101 (Texas Instruments) are solely solid-state devices using an electroluminescent junction diode irradiating a junction-type photoelectric semiconductor device such as a PN or PIN diode, a phototransistor or a 4-layer diode (Fig. 2-59C). These are fast-response devices featuring rise times of from 5 ns to 15 microseconds, according to type, and cutoff frequencies of 2 to 8 MHz. They can stand voltages of up to 5,000 applied between input and output circuits and even up to 50,000v using fiber optics. The input-output capacity is small, however.

This directional optical link makes easier the solution of a lot of input-output loop problems where ordinary quadripoles are often difficult to use. For **on-off type controls**, Fig. 2-59 suggests a relay function standing practically unlimited voltages between input and output circuits. Fig. 2-60 illustrates the operation of a **chopper** whose applications are discussed later. The low DC transducer voltage e is applied to points A and B of a voltage divider composed of photoconductor Pc and resistor R. Pc is illuminated by a glow tube. Suppose, now, the Pc dark resistance of 10 meg is lowered to

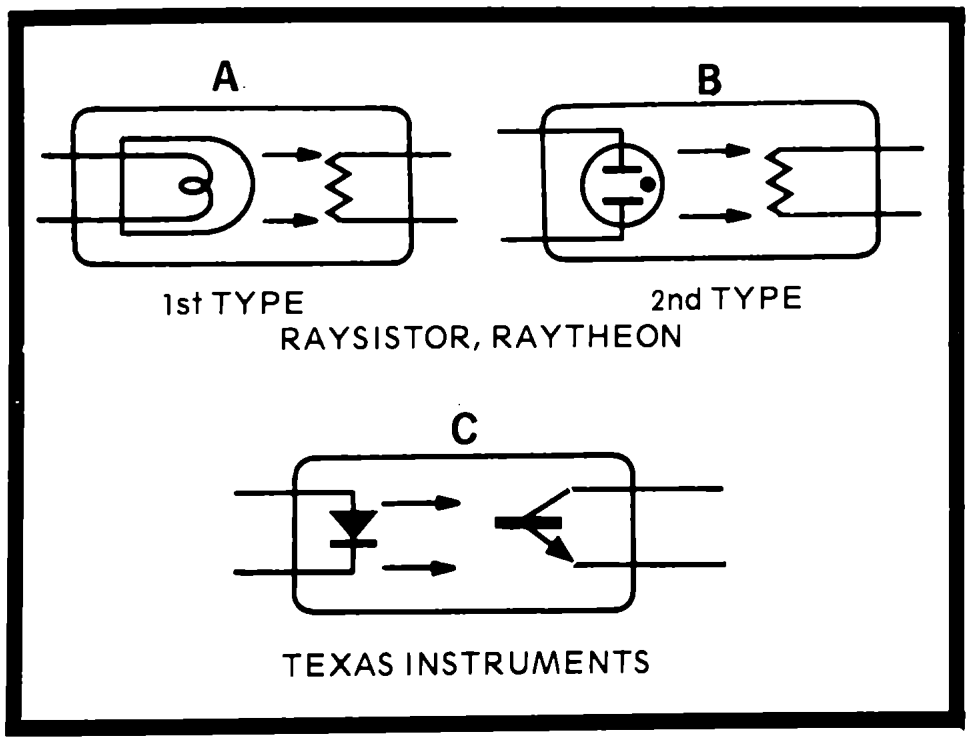


Fig. 2-59. Schematics of three different types of photoelectric choppers.

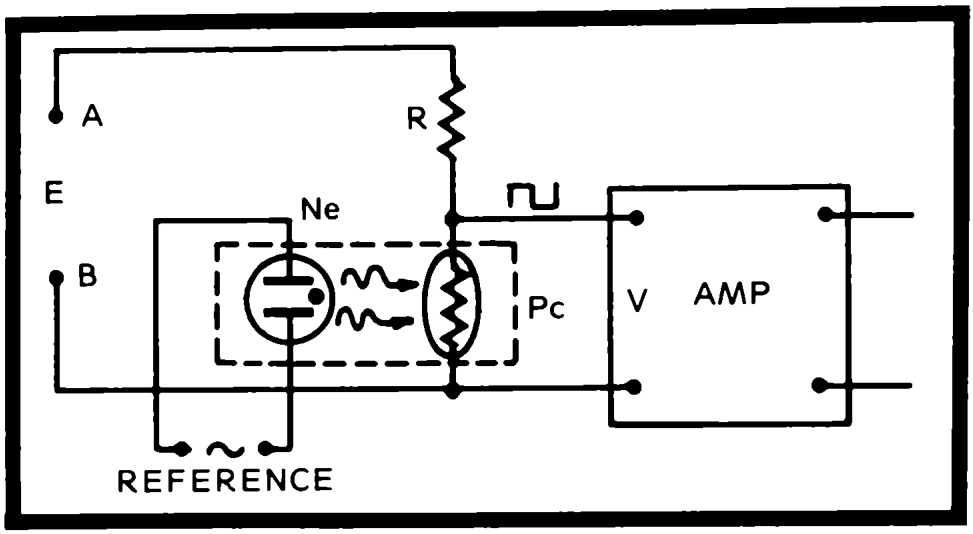


Fig. 2-60. Photoelectric chopper for low-level DC voltages.

1K upon illumination. If R is 0.1 meg, amplifier input voltage v will be a square wave with top and bottom levels of about e divided by 100 and e; the switching of the glow tube is accomplished by the AC reference voltage.

The device in Fig. 2-59A is also serviceable as a low-frequency chopper. The thermal inertia of the tiny lamp filaments used is sufficiently low to get a sufficient brightness fluctuation of about 5 percent at 60 Hz. Connecting diode D in series with the lamp makes for lighting it only 60 instead of 120 times per second (Fig. 2-61). The push-pull chopper in Fig. 2-62 uses two alternatingly operating light modulators featuring an efficiency approaching 100 percent for frequencies of from 10 to 100 Hz. For the more recent solid-state choppers, the cutoff frequency is so high that it can be ignored for control purposes.

These modulators also allow for the economic design of logic gates (discussed in Chapter 5); the electrical isolation

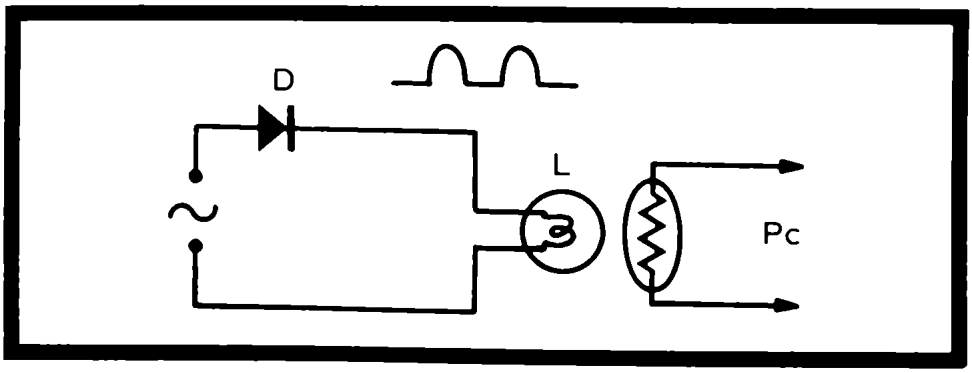


Fig. 2-61. The rectifier diode reduces the lamp flashing rate from two to one per cycle.

between input and output due to optical coupling may be a precious advantage here. The OR, AND and NOT gates in Fig. 2-63 compare with those in Figs. 5-1, 5-2 and 5-3. It is easily seen that to light the lamp, either Pc1 OR Pc2 of the OR gate must be illuminated, while the AND gate requires simultaneous illumination of both Pc1 AND Pc2.

Between the applications of the optic modulator to continuous controls (working by the more-or-less rather than the on-off principle), there are regulators of electrical or other magnitudes. Because of its discontinuous operating characteristic, the glow tube cannot be used here. In the circuit in Fig. 2-64, in a manner somewhat similar to inverse feedback, the input of an amplifier is applied by means of a voltage divider in which photoconductor Pc is a parallel resistive element; the amplifier input voltage, therefore, decreases with an increasing illumination of Pc. If Pc is illuminated by lamp Ls, connected across the amplifier output, we are confronted with a voltage stabilizer. Lamp Ls, series connected with load RL, tends to stabilize the output current.

The circuit in Fig. 2-65 allows for stabilizing the brightness level of lamp L, DC-fed by a Darlington circuit composed of current amplifier transistor Q1 and controlled series transistor Q2. Photoconductor Pc, connected in the divider supplying the base of Q1, is illuminated by L. A brightness increase results in lowering the resistance of Pc, during which time the base of Q1 is going more positive; the resistance of Q2

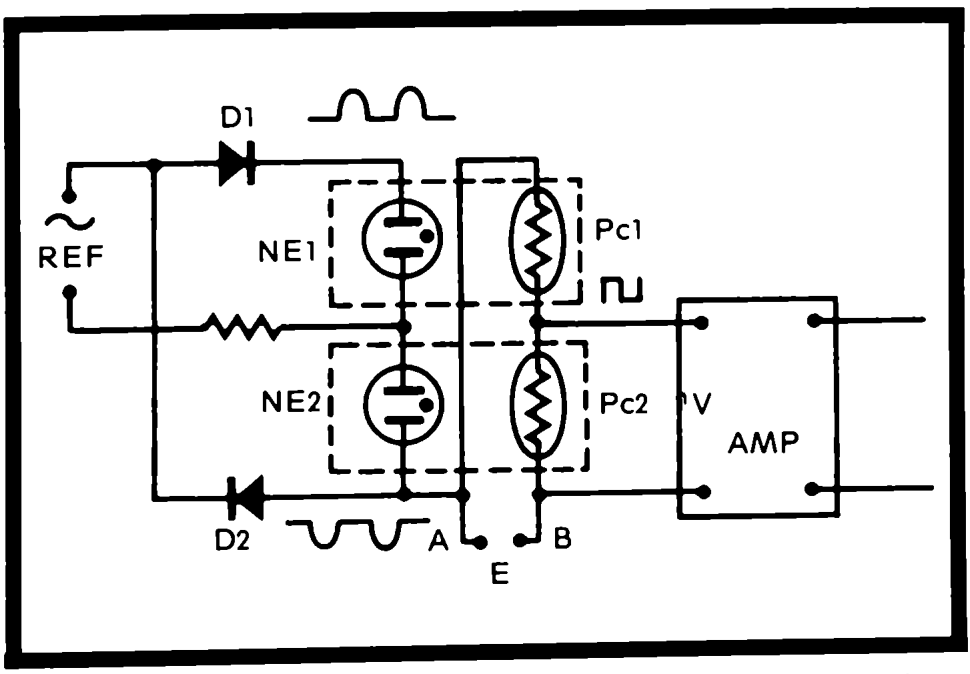


Fig. 2-62. The push-pull chopper has better efficiency than the chopper in Fig. 2-61.

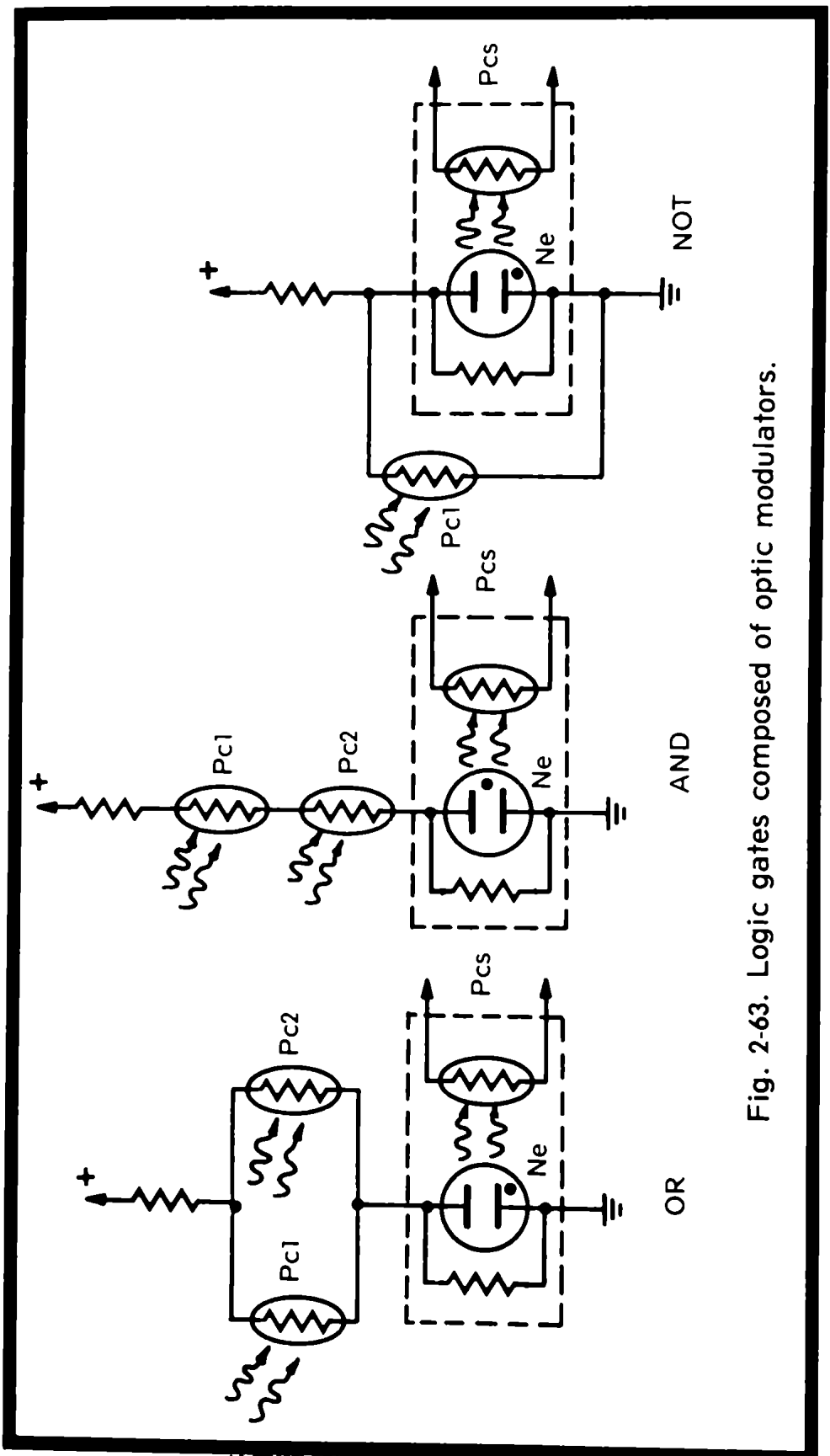


Fig. 2-63. Logic gates composed of optic modulators.

UNIVERSITY OF VICTORIA

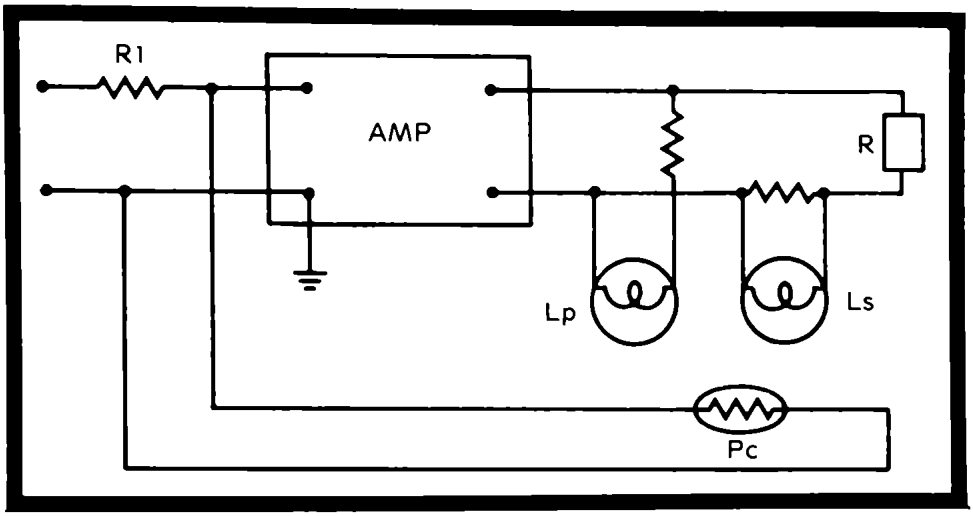


Fig. 2-64. Diagram of an optic modulator wired in an inverse-feedback loop.

increases, thereby lowering the brightness toward its original value. The regulation level can be set electrically, by varying R1 or R2, or optically by varying distance **d** between L and Pc, or by inserting filters or a diaphragm in the light path.

HALL-EFFECT TRANSDUCERS

Although the Hall effect was described for the first time about one hundred years ago, it had to await the development of semiconductors to be of practical use, for it is insufficiently

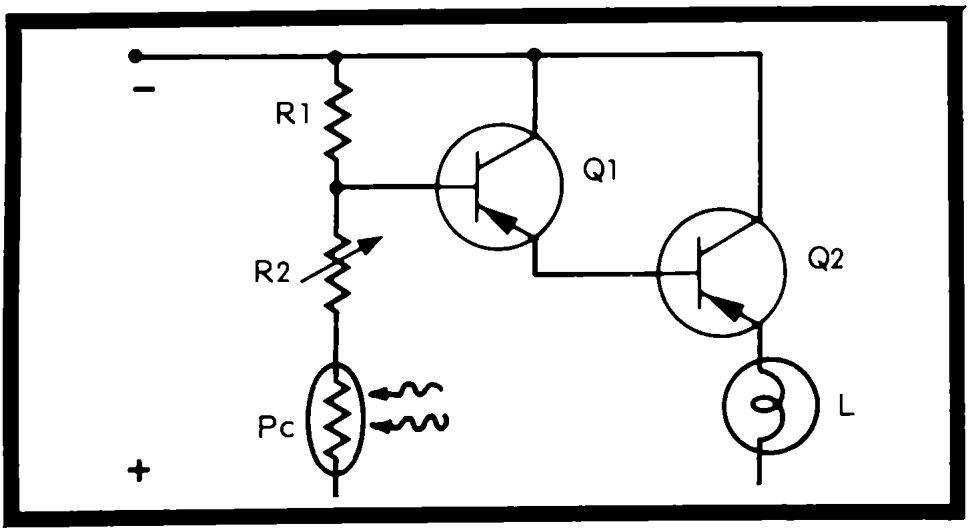


Fig. 2-65. This optic modulator feedback circuit regulates the brightness level.

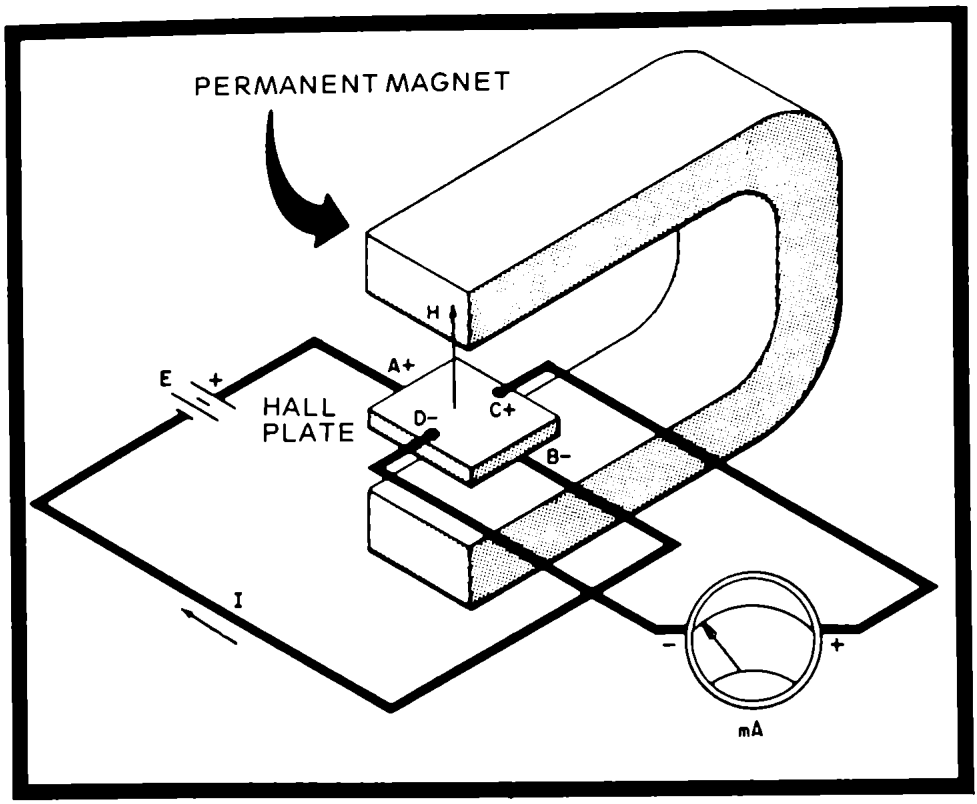


Fig. 2-66. Drawing showing the arrangement of the semiconductor platelet in a Hall effect device.

pronounced in good conductors (metals). A Hall-effect transducer is made up of a rectangular slab of semiconductor material bearing two pairs of facing electrodes A+ and B-, C+ and D- (Fig. 2-66). A battery passes a direct current between A and B, and a magnet M sets up a magnetic field H in the air gap containing the slab. Current I and field H induce a DC (Hall) voltage VH between C+ and D- so that VH is kIH, where k is a constant depending upon the material used and its physical dimensions.

In regard to constant current I, VH is proportional to the magnetic field H; this is a very convenient method for measuring magnetic fields and related magnitudes. Such transducers can be made very thin, allowing for the exploration of magnetic fields even within small air gaps. Alternately, with a magnetic field set up by a current to be measured, the device becomes a wattmeter for DC as well as for AC (Fig. 2-67). In the case of AC, the device takes phase angle into account and yields the real power EJ cos ϕ. This phase sensitivity is the principle underlying Hall effect phasemeters.

In analog computers, Hall-effect transducers are used as multipliers, for if H and I are made proportional to the

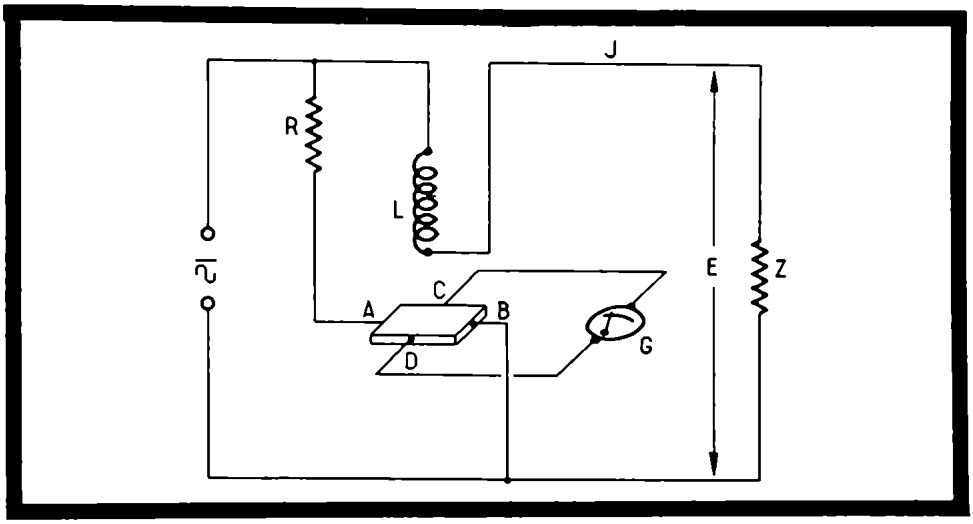


Fig. 2-67. Wattmeter circuit using a Hall effect transducer.

variables, the device carries out their analog multiplication, since the Hall voltage VH is proportional to the product IH.

A very sensitive Hall-effect **displacement transducer** is schematically represented in Fig. 2-68. It is composed of a fixed magnetic C core cut in its center to allow insertion of the Hall plate and a small mobile magnet moving just above the

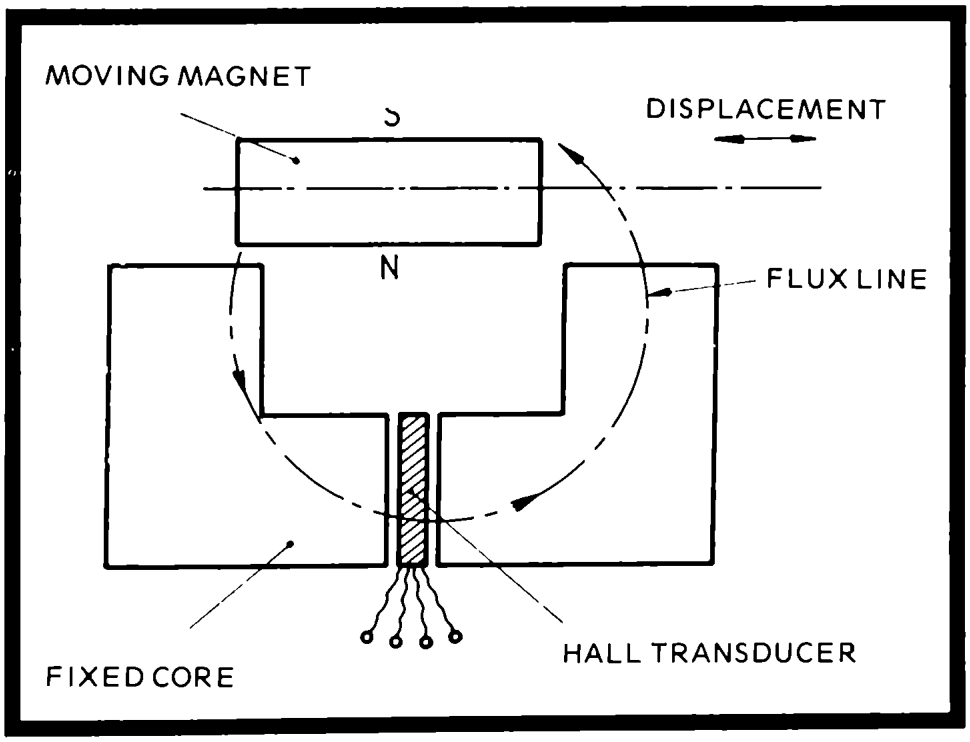


Fig. 2-68. Diagram of a Hall effect displacement transducer.

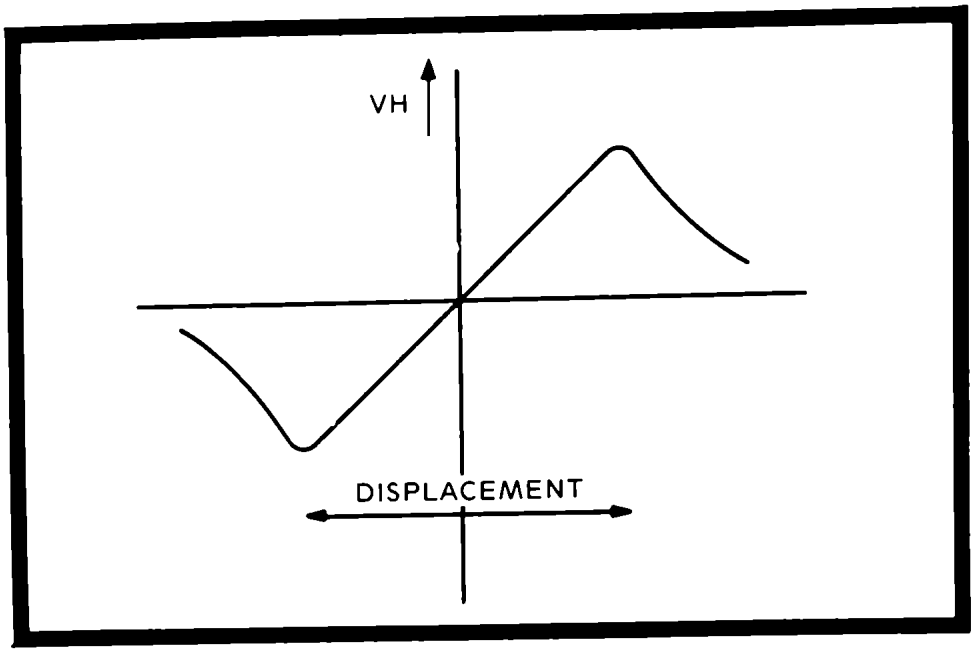


Fig. 2-69. S characteristic curve of a Hall effect displacement transducer.

large air gap, mechanically linked to the mobile part whose displacement is to be monitored. Provided a constant current I is supplied to electrodes A and B, voltage VH variable in terms of the magnet position, develops across C and D. As the magnet is polarized vertically, the flux lines leave the magnetic north on the lower face, pass through the iron core and re-enter the magnet by the south on the upper side. Disliking air gaps, the flux choses a path with the least possible air gap length. If the magnet is magnetically centered in regard to the iron core, there is no flux, and the VH versus displacement characteristic crosses the zero line (VH equals 0). As the flux enters the iron core by the leg nearest to the off-centered magnet, the polarity of VH indicates the direction of displacement, and a characteristic such as shown in Fig. 2-69, similar to that of a frequency or phase discriminator, results. The device is quite linear as shown between the specified displacement limits.

Just like the extensometer manufacturers who have devised numerous methods for measuring a lot of magnitudes by means of strain gauges, Hall-effect transducer manufacturers (especially German) have developed techniques using such devices to measure more and more magnitudes, and technical papers frequently describe new developments. A laboratory mastering a given technology with equipment of its proper design will evidently find it more profitable to find new applications for methods and material on hand than to research and develop competitive techniques.

CHAPTER 3

System Building Blocks

All electronic systems are composed of basic devices such as amplifiers, oscillators, detectors and others, forming "building blocks" interconnected to get the desired performance of the whole. Of course, it is impossible to treat exhaustively in a few pages so wide a subject. Thus, only basic principles are outlined here, emphasizing typical electronic control circuits.

AMPLIFIER TYPES AND CHARACTERISTICS

To begin with, consider the "black box" labeled amplifier (Fig. 3-1) presenting two input posts **a**, **b** and two output posts **c**, **d**. (Posts **b** and **d** are conveniently tied together and grounded whenever possible.) Applying an input voltage or signal Vi at the input results in an output voltage Vo at the amplifier output. Both voltages are of similar waveform, but their amplitude and phase may differ. (Control amplifiers generally do not require the perfect linearity characteristic sported by hi-fi equipment, especially so if they are pulse operated.)

If Vo is greater than Vi, there is voltage amplification, and the actual gain is:

$$G = V_o / V_i$$

Though less evident, there may still be amplification if Vo is less than Vi. Suppose the black box input resistance, Ri is 1 meg and load RL is 100 ohms; Vi is 100v and Vo is 10v. While the voltage gain of 0.1 actually is a loss, the input and output powers, respectively, are

$$P_i = V_i^2 / R_i = 0.01w$$

and

$$P_o = V_o^2 / R_L = 1w$$

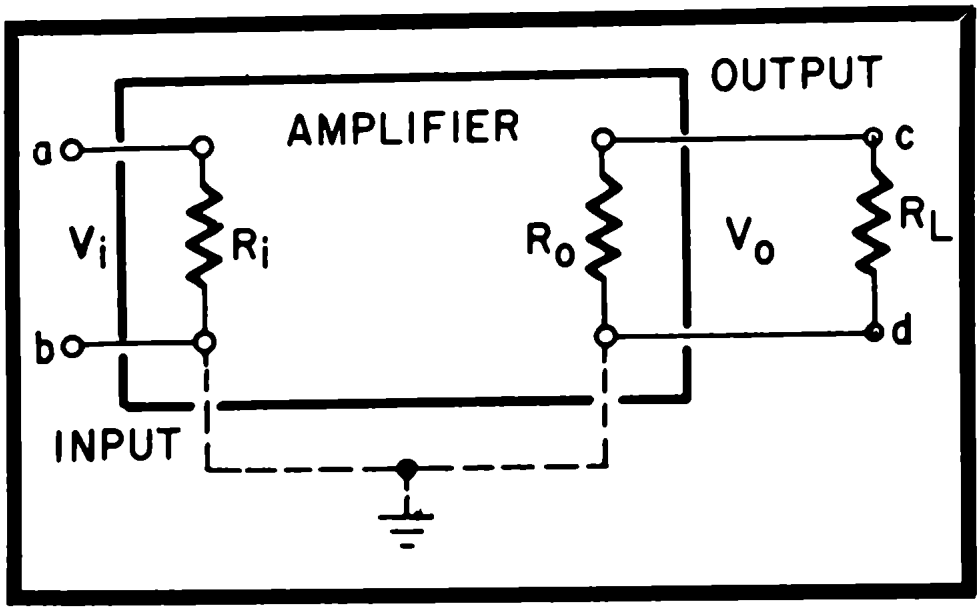


Fig. 3-1. This diagram illustrates the main parameters of an amplifier.

Hence, there is a **power gain** of:

$$P_o / P_i = 100$$

Similarly, a **current gain** of 1,000 is obtained.

Power amplifiers (one stage often will do) provide the power necessary to energize motors, electromagnets and other kinds of loads. A voltage or current **driver amplifier** generally is required to sufficiently amplify the low-level signal for controlling a power amplifier. The gain required determines the number of driver amplifier stages needed.

The transmitted frequency band is another fundamental amplifier characteristic. In industrial electronics, mostly low and very low frequencies are used, and 100 kHz can be considered as an upper limit. (This does not include special HF gear not "interfacing" with conventional control equipment.) On the other hand, phase distortion-free amplification of signal frequencies lower than 60 Hz is not too easy, for the use of RC-coupled AC amplifiers may lead to excessive time delay and instability. There are so-called DC amplifiers (i.e., amplifiers for slowly varying voltages or currents) that are either direct-coupled or chopper amplifiers.

Amplifiers contain electron tubes or transistors somewhat modulating the DC supply voltage or current to provide an enhanced amplitude output signal. Although both devices accomplish the same functions, they do it differently, and, therefore, cannot be substituted.

INVERSE FEEDBACK

Consider the black-box type of amplifier in Fig. 3-2. It produces an output voltage, Vo, upon applying an input voltage Vi. By means of output-fed voltage divider R1, R2, a portion β of Vo is reinjected into the output ("feedback") in a way to oppose the initial signal Vs (inverse feedback). Without any feedback, the amplifier gain was:

$$G = Vo/Vi$$

Amplifier input Vi is identical with signal input Vs. By injecting feedback voltage βVo, in series with Vi, the overall gain reduces to:

$$G_o = V_o/V_s = V_o/(V_i - \beta V_o)$$

Hence:

$$G_o = G/(1-\beta G)$$

Assuming the factor β G is far greater than unity, this reduces to 1 divided by β . This means that the overall gain of that amplifier practically depends on:

$$\beta = R_2/(R_1 + R_2)$$

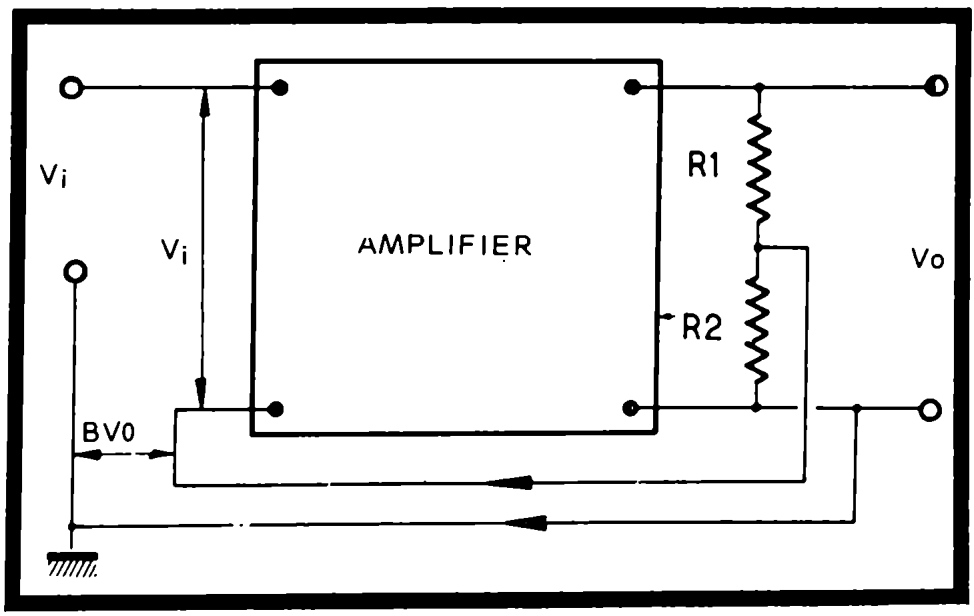


Fig. 3-2. Amplifier with an inverse voltage feedback loop.

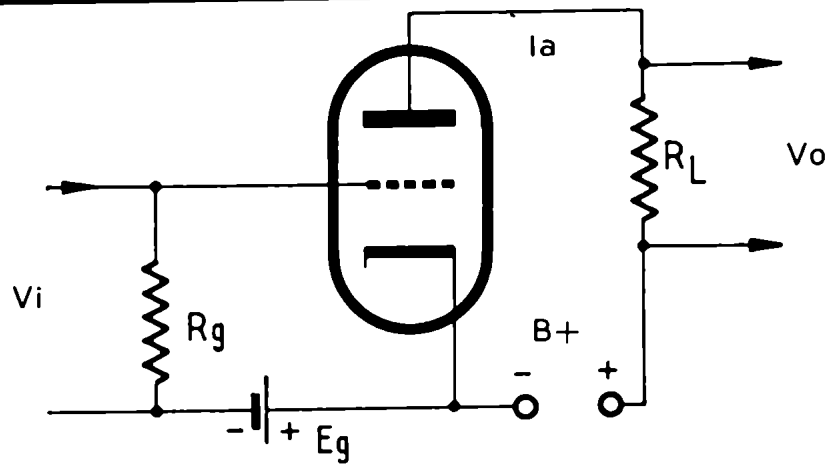


Fig. 3-3. Schematic of a basic triode amplifier stage.

and is independent of its original gain G. Supply voltage variations, aging and other factors do not influence Go. Using a high feedback factor, the apparently varying device becomes a stable and accurate amplification standard, at the price, however, of a somewhat lower gain.

VACUUM TUBE AMPLIFIERS

Now let's look inside the black box. A triode amplifier is represented schematically in Fig. 3-3. The anode is made positive regarding the cathode by a B+ power supply generally providing from 100 to 300v, and the grid is negatively biased by a suitable source delivering a few volts under negligible current. (Remember, there is also a heater to be energized, for the cathode will emit the required electrons only when heated to about 1,600 degrees F.) The electrons set free by the cathode are attracted by the (positive) anode, and an anode current, Ia, passes through load resistor RL. To reach the anode, the electron flow has to pass the grid, but the negative grid bias repels the electrons and thus decreases (and even cuts off) Ia. This principle allows for control of Ia—and the voltage drop (Vo) across RL by the same token—by a mere variation of the grid potential requiring only a negligible amount of power.

To obtain a higher gain, another grid, called the screen, has been inserted between the control grid and anode. This screen is made positive by resistor Rs tied to B+; it is held, AC-wise, at cathode potential by bypass capacitor C (Fig. 3-4). The screen accelerates the electrons issued by the cathode independently of the anode potential, thereby strongly in-

creasing the plate resistance of the tube. As the mutual conductance remains unchanged, K reaches a fairly high value, 1,000 and up. The screen also cancels the undesirable capacitive coupling between input and output circuits, particularly inconvenient in HF amplifiers.

The tetrode tube thus obtained presents an N-shaped Ig over Va characteristic; if Va is smaller than Vs, the mutual conductance becomes negative, and the amplifier becomes unstable at least. This inconvenience is eliminated by inserting a third grid, called a **suppressor**, between the screen grid and anode. Tied to the cathode, the suppressor repels secondary anode electrons normally attracted by the more positive screen grid. This 5-electrode tube, appropriately called a **pentode**, is represented in Fig. 3-4, showing electrode arrangement and connection.

DIRECT-COUPLED ELECTRON TUBE AMPLIFIERS

As electron tubes are superseded more and more by transistors in amplifier design, available space is devoted mostly to solid-state techniques; however, there are a good many textbooks dealing with conventional tube amplifiers to be consulted by interested readers.

Although it is relatively easy to cascade direct-coupled transistor amplifiers because of the low supply voltages in-

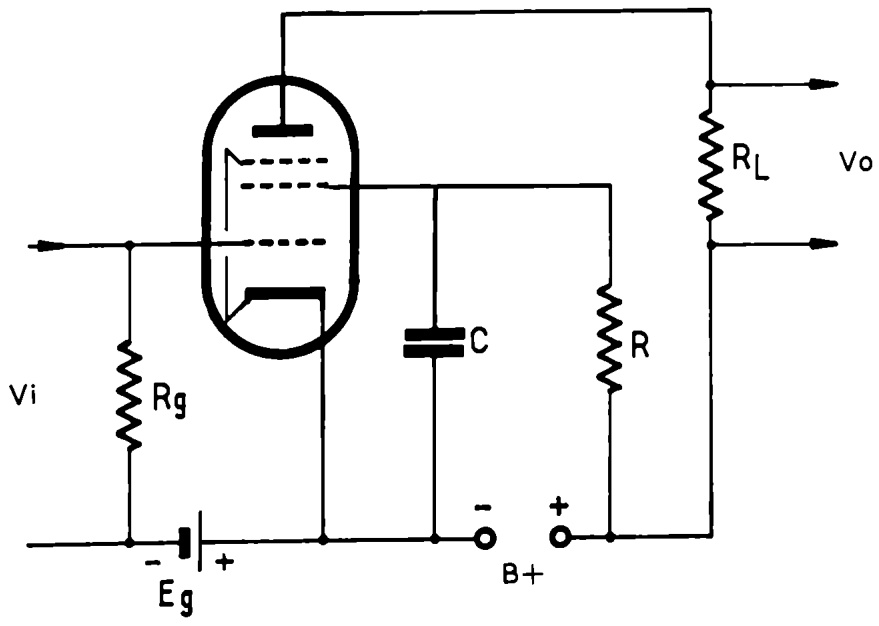


Fig. 3-4. Basic pentode tube circuit.

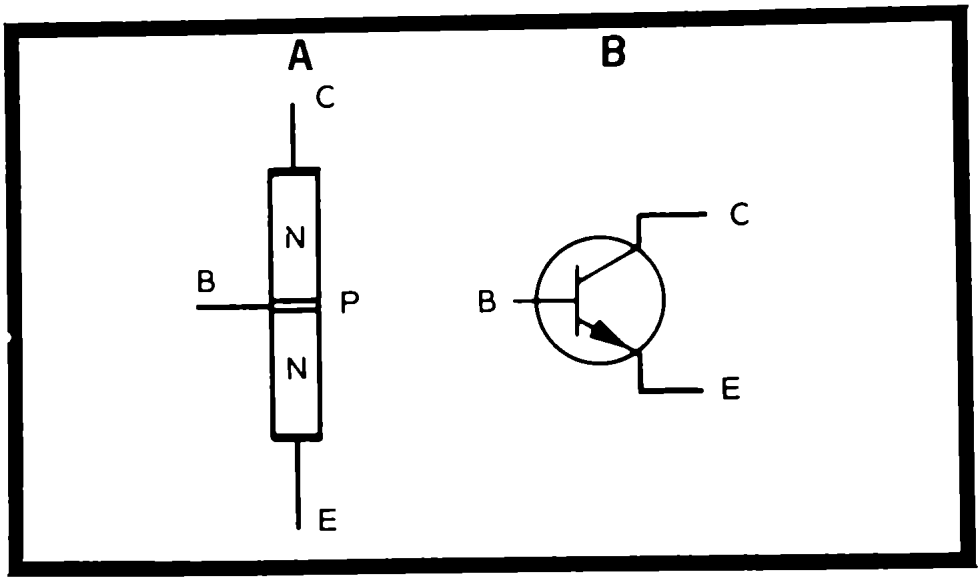


Fig. 3-5. Structural diagram and schematic symbol of an NPN transistor.

volved, the same problem is more delicate to solve using tubes, for with a grid-anode voltage difference generally somewhere between 50 and 250v, a very high B+ supply inconvenient to use results. not to speak of the multiple heater windings needed because of the high voltage differences between the heaters of such cascaded tubes.

THE TRANSISTOR

Like the electron tube, the transistor is an active device, and it is able to amplify. Notwithstanding this fundamental similarity, its operation is based on a quite different principle; therefore, a tube cannot be replaced directly by a transistor. A transistor (NPN-type junction) is shown cut lengthwise in Fig. 3-5A, while Fig. 3-5B indicates the corresponding symbol. The tiny semiconductor bar, generally made of germanium or silicon, is shown much enlarged here. By introducing an impurity such as arsenic, antimony or phosphorus, the non-conducting pure semiconductor can be made conducting by (negative) electrons. In the center of the bar has been arranged a small P region of positive hole conduction, the holes being vacancies left by the departure of electrons.

The electrodes placed at the ends of the bar are called emitter (E) and collector (C), and the center electrode is the base (B). The name and symbol of the base are reminiscent of the early transistors where two point contacts (cat whiskers) touched a semiconductor chip used as a base.

Because of its two junctions NP and PN, this transistor behaves just like two opposing diodes with a common cathode, B (Fig. 3-6). With E grounded, let us apply positive potentials to B and C and look at the meter scales to see what happens. If B were unconnected, collector current Ic would be zero whatever the polarity of collector battery B1, for if one of the diodes were conducting, the other would be reverse biased and block any current. For the polarity shown, the upper diode (BC) is cutoff. If now we make B positive regarding E by means of potentiometer P, shunting battery B2, the lower diode (EB) is thereby biased to conduct and will pass the electrons provided by emitter E (hence, its name), and since the base region is made extremely thin by design, the electrons attracted by the positive collector potential will easily traverse the base to penetrate the collector region. The collector thus absorbs practically from 95 to 99 percent of the emitter current, with only 1 to 5 percent of it taken by the base.

Considering that the electron tube requires a high voltage B+ as well as heater supplies, the transistor offers an amazing flexibility in circuit design. Each of its three electrodes can be made the common input-output electrode as deemed convenient, resulting in the three basic configurations—**common emitter** (CE), **common collector** (CC) and **common base** (CB) diagrammed in Fig. 3-7. Most amplifying transistors are CE-connected, yielding an input-output current gain of 20 and more. While tubes have a very high input, and a medium to low output resistance, the transistor input resistance is lower than its output resistance, making for high voltage and power gains. The voltage gain of the CC configuration is a little less than one, but there actually is current and power gain; this connection is used as a step-

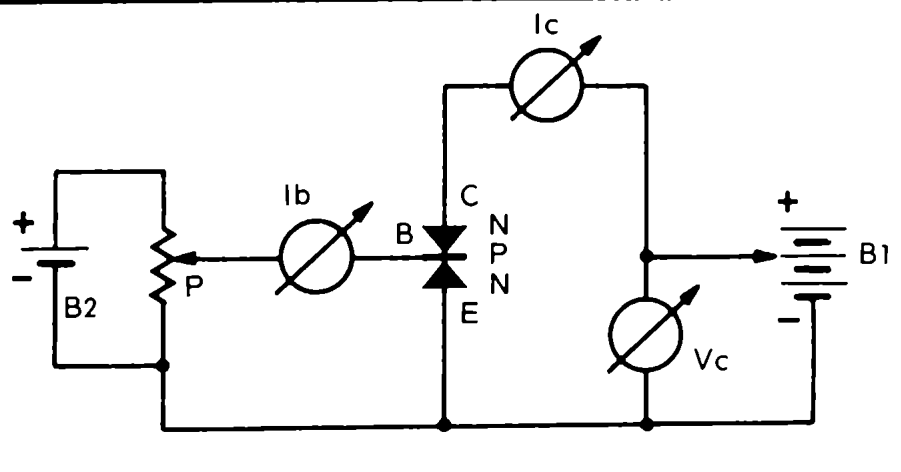


Fig. 3-6. A transistor behaves like two opposing diodes.

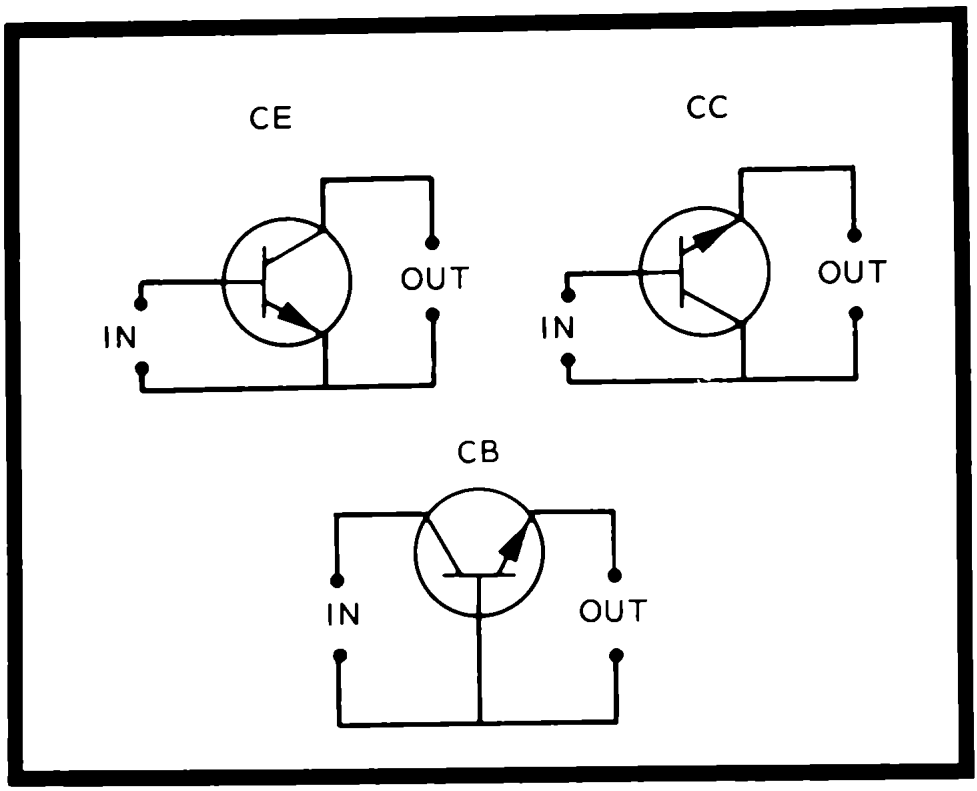


Fig. 3-7. Schematics of the three basic transistor configurations.

down impedance converter, especially for low input impedance CE amplifier transistors. The CB connection offers no current gain because of its low input and high output resistances; a voltage or power gain is possible. This configuration is principally used in high-frequency oscillators.

On the other hand, if instead of the N-type bar with a P region, we take a P-type bar with an N region, we get a PNP transistor complementary to the former NPN type (Fig. 3-8). With similar characteristics provided, the PNP transistor behaves exactly as the NPN type does, except for the fact that the polarities of all voltage sources used are reversed. This singular property, which has no counterpart in electron tube technology, based on electron rather than hole conduction, allows for interesting possibilities in circuit design, since it offers additional flexibility.

TRANSISTOR AMPLIFIERS

The collector current (Ic) of a transistor is controlled by its emitter-base current Ieb. This makes for a low input resistance—a few hundreds to some thousands of ohms. The preceding stage has to match this load. If it is a similar

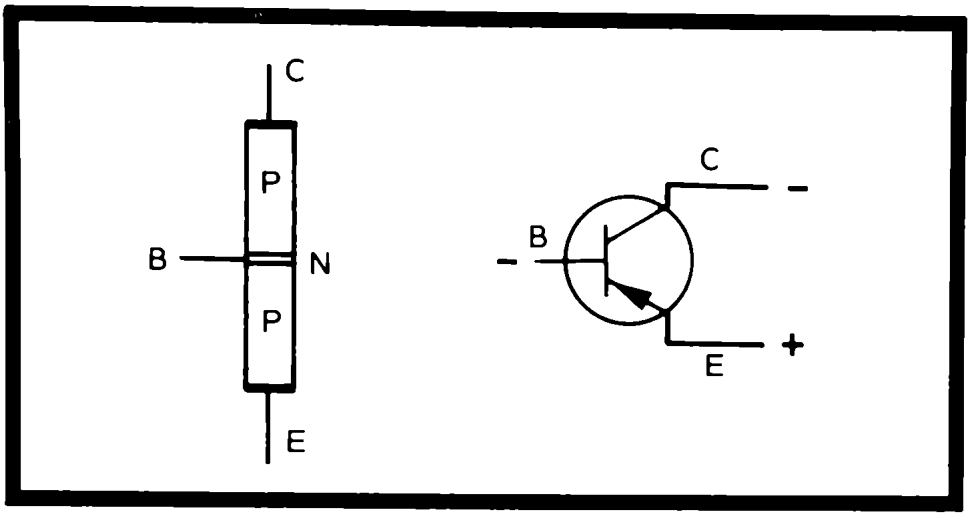


Fig. 3-8. Structural diagram and schematic symbol of a PNP transistor.

transistor stage, its output resistance is comparatively high, 50,000 ohms, for instance. There evidently is a problem.

On the other hand, the emitter-base junction of resistance re is common to base and collector circuits. This results in an internal feedback between input and output, greatly complicating the calculation of transistor circuitry. A number of equivalent quadripoles describe transistor operation in a given set of conditions.

The simplest possible transistor amplifier is represented in Fig. 3-9. A PNP transistor is CE-connected; for an NPN, reverse the supply battery as well as the polarity of the electrolytic capacitors. The base is biased to the same polarity as the collector by series resistor Rb. To obtain a base bias current of 0.1 ma from a 6v battery, Rb is 60K. (Regarding Rb, r_e is negligibly small.) The output voltage develops across collector resistor Rc. Coupling capacitors are provided to avoid shorting out the electrode voltages.

This simple biasing method is seldom used. Transistors are marketed with some allowance on the current amplification factor β, so that Rb should be matched to the transistor used. Still more important, the internal transistor resistances are largely temperature-variable and may lead to thermal runaway and eventually to destruction of the device. Consequently, it is necessary to choose a bias method that compensates for thermal effects. In the amplifier stage in Fig. 3-10, an excessive base current is provided by voltage divider R1R2, and an opposing bias is developed by emitter resistor Re, bypassed by a suitable capacitor (C) to avoid a loss of AC signal components. This circuit yields a gain of about 172; a high load impedance, however, is required to make use of it.

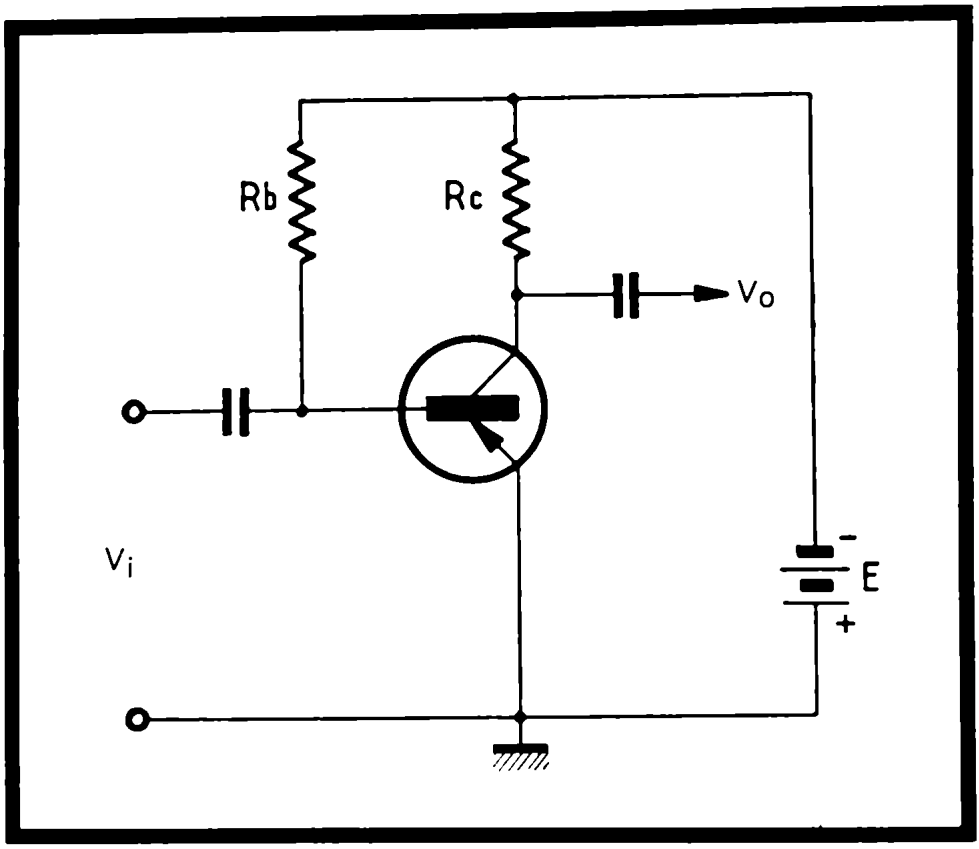


Fig. 3-9. Basic common-emitter (CE) transistor amplifier circuit.

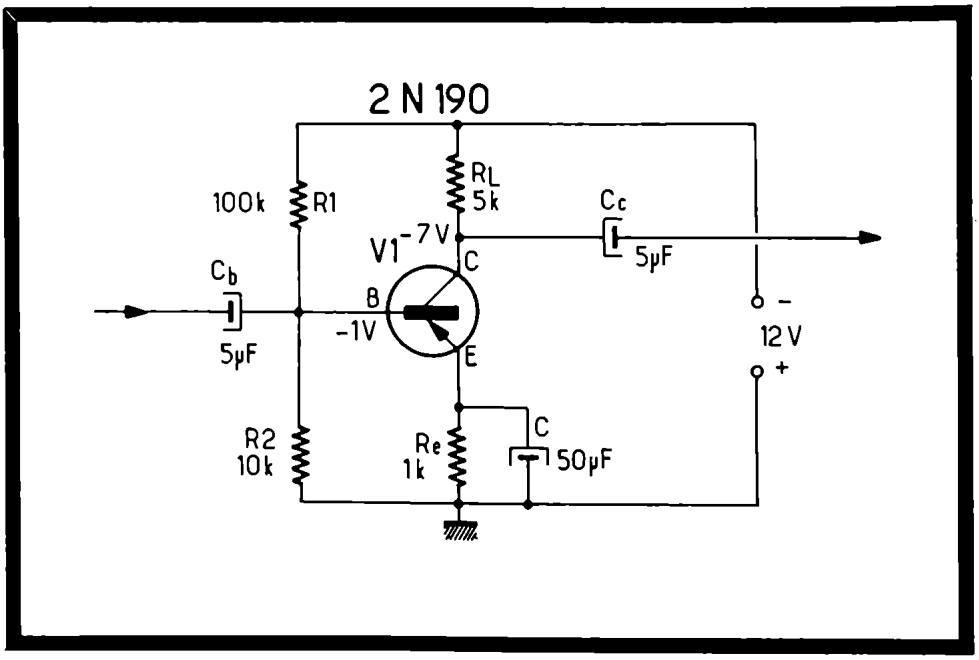


Fig. 3-10. CE amplifier with thermal stabilization.

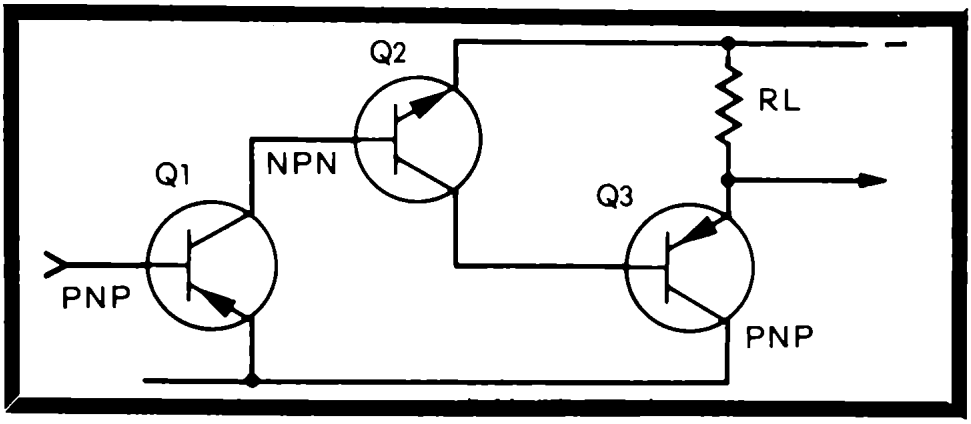


Fig. 3-11. Direct-coupled 3-stage amplifier using complementary transistors. Only one passive component is used (RL).

To get a higher gain, a number of stages of this type can be cascaded. Taking advantage of the peculiar proprieties of transistors, the 3-stage amplifier using complementary transistors in Fig. 3-11 uses the fewest possible passive components (only one resistor and no capacitor) and thus is a natural for microelectronics. The base currents of Q2 and Q3 equal the collector current of the preceding transistor, and the overall current gain equals the product of the individual gains of the transistors used. The voltage gain of this amplifier is given by the product of the mutual conductance of Q1 the current gain of Q2 and Q3 and the value of RL.

While a power tube is driven by a **voltage amplifier**, power transistors require a **current amplifier** drive. The so-called **Darlington** connection is frequently used for this purpose. As shown in Fig. 3-12, CC stage Q1 directly drives, from the emitter, the base of CE-connected output stage Q2. The overall

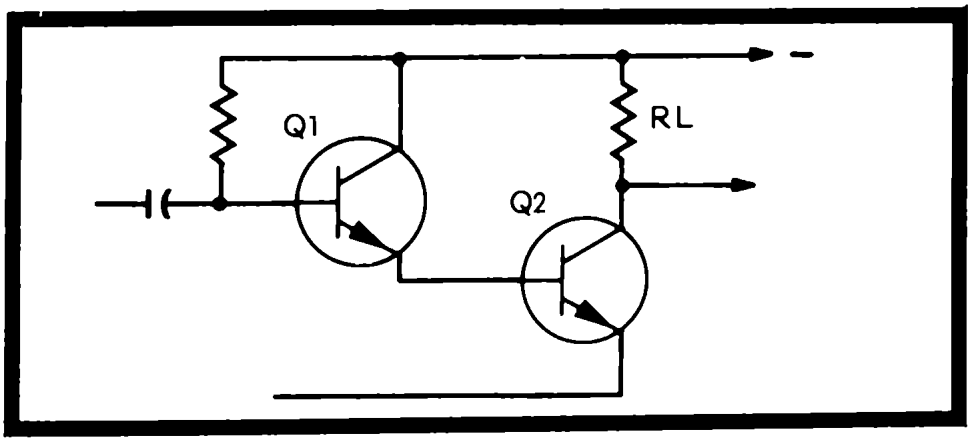


Fig. 3-12. Schematic of a Darlington-connected current amplifier.

current gain, again, equals the product of the current gains of Q1 and Q2, while the voltage gain of Q2 makes up the entire voltage gain. The input resistance is that of Q2 multiplied by the current gain of Q1, an increase to be welcomed.

Unlike a tube, the transistor will directly (transformerless) drive a low-impedance load (several ohms).

THERMAL COMPENSATION OF DIRECT-COUPLED TRANSISTOR AMPLIFIERS

Because of the low supply voltages involved, generally between 1.5 and 24v, direct-coupled amplifiers are easier to design with transistors than with tubes. But on the other hand, the effects of thermal variations in transistors (more pronounced in germanium than in silicon) may well add and lead to a thermal runaway unknown in tube circuits. Therefore, it is determined first if the thermal effects tend to accumulate or cancel. Thermistors and reverse-biased diodes, in which the resistance varies with temperature, are often used to compensate for thermal effects.

The two direct-coupled CE amplifier stages in Fig. 3-13 feature a built-in thermal compensation tendency, for an increase of T1 collector current results in a lower collector-emitter voltage of the same transistor, thereby lowering the base-emitter voltage of T2, counteracting the increase in its collector current. Diode D, used as load for T1, still improves this compensation by its negative temperature coefficient.

FIELD-EFFECT TRANSISTORS (FET)

We know that tubes are voltage-controlled and transistors are current-operated. This is true for conventional, so-called **bipolar** transistors, but not for **unipolar** or **field-effect** transistors (FET). A junction-type FET is composed of an N-type silicon chip with electrodes on the two opposite ends, called **source** S and **drain** D. On the central part of this chip, P-type regions are diffused into the upper and lower sides; metallic contacts are made and tied together to make up the **gate** (G) electrode (Fig. 3-14). If the drain-source resistance of the device is connected across battery B1 through resistor RL (Fig. 3-15), there appears (proper polarity provided) a **depletion region** extending more toward the center of the chip with an increasingly higher negative gate potential referred to source. The results is a "pinch" of the drain-source current, due to the electrostatic gate field. The gate, therefore, acts to control current in a way somewhat similar to the grid control

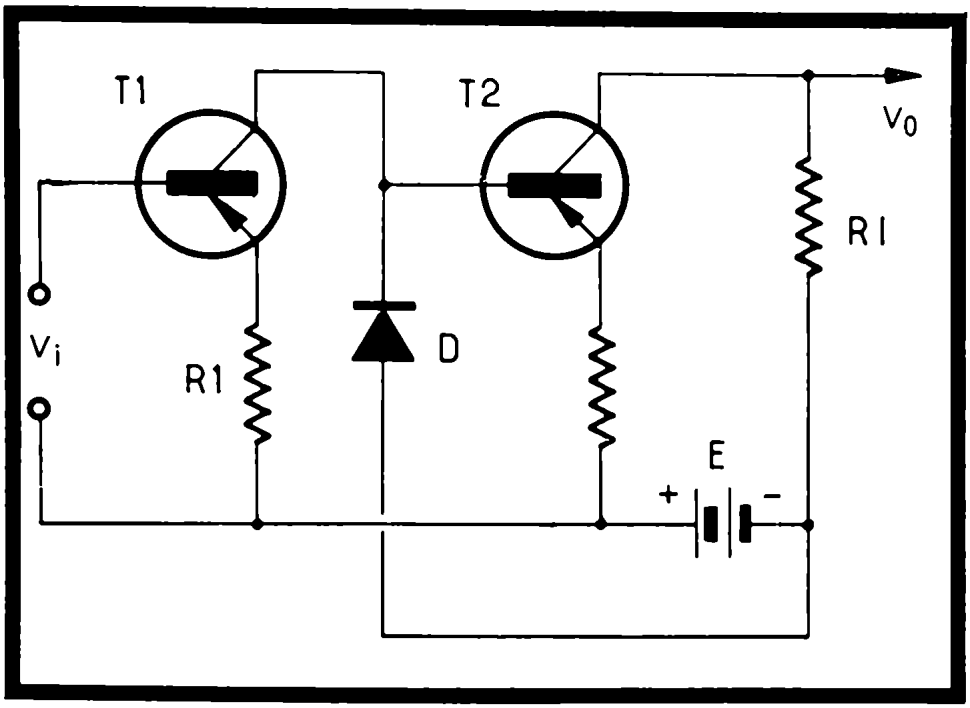


Fig. 3-13. Direct-coupled transistor amplifier with improved thermal compensation.

of an electron tube, though the operating principles are entirely different.

The likeness is even more disconcerting when examining the drain current in terms of the drain-source voltage (L_D over V_{DS}) characteristic for a set of gate potentials UG; these

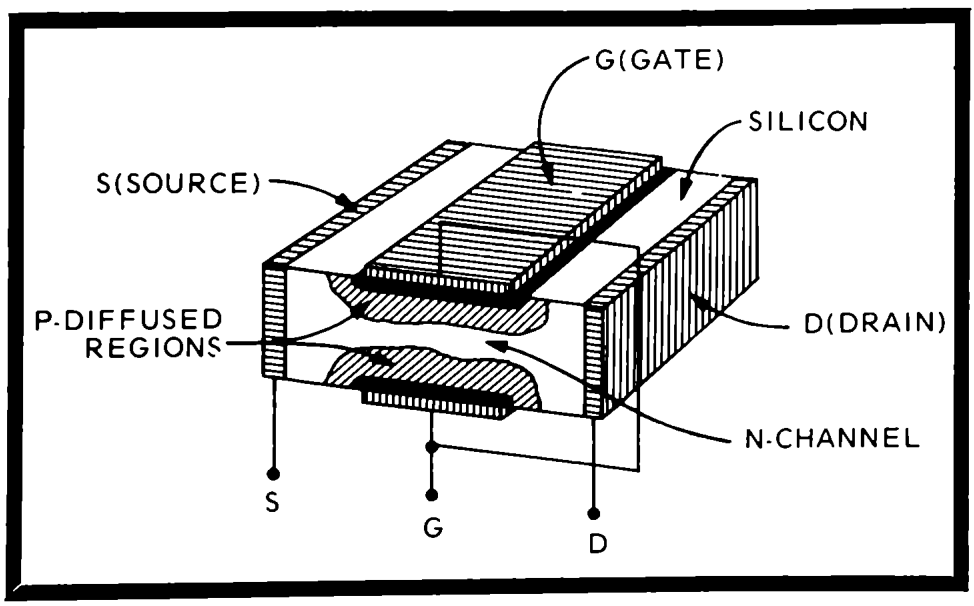


Fig. 3-14. Structural diagram of a junction-type field-effect transistor.

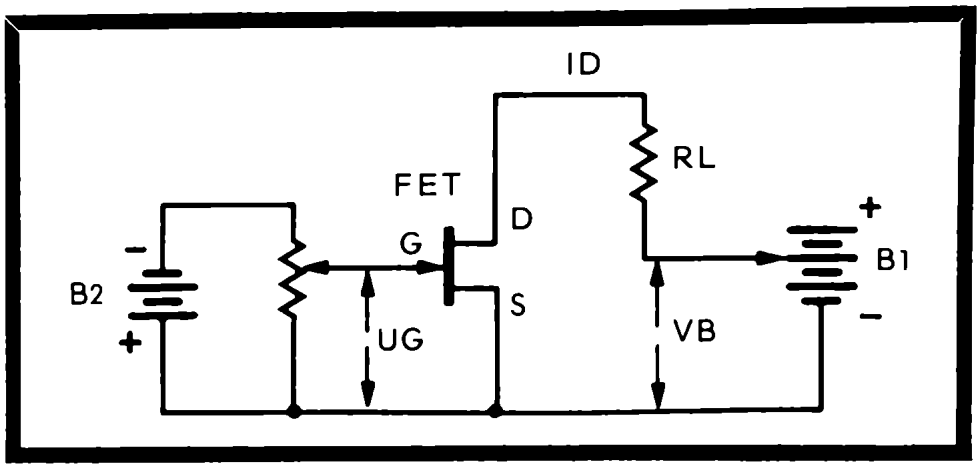


Fig. 3-15. Basic FET circuit.

curves (Fig. 3-16) are reminiscent of the Ia over Va characteristics of a pentode. The gate forms, with the chip material, a reverse-biased PN junction, and the input impedance of the FET is at least as great as that of a pentode tube. For a number of models, manufacturers specify a **typical** leakage current of 1na (one nanoampere, or 10^{-9} ampere)! Therefore, FETs are a natural for various circuits requiring a very high input impedance, far exceeding the possibilities of con-

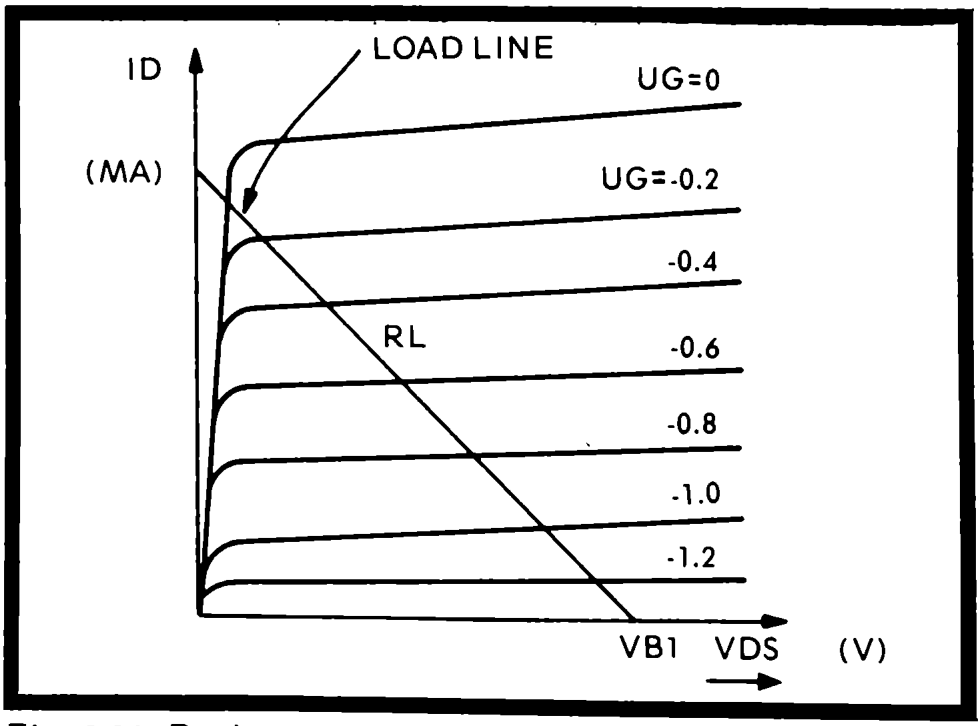


Fig. 3-16. Drain current (ID) drain-source voltage (VDS) characteristics of a typical FET for various gate voltages.

ventional semiconductor devices. Remember, however, that this junction is nothing more than a reverse-biased silicon diode, and that it is serviceable only at low voltage levels.

The MOSFET gate acts by its electrostatic field through the dielectric on the space between the two N regions; its control action is strongest with the thinnest dielectric film, and a comparatively small potential difference may be enough to puncture this layer, resulting in device destruction. Particular measures, therefore, are to be taken to protect the MOSFET against any voltage transients and static charges a junction FET will stand. Thanks to this dielectric, the MOSFET input resistance is even higher than that of the junction FET—from 1012 to 1015 ohms against about 1010 ohms, this leakage actually is that of the insulated feedthroughs of the sealed casing. This dielectric layer between the gate electrode and the semiconductor material is depicted in the MOSFET symbol.

If the MOSFET is connected like the FET in Fig. 3-15, an N channel (called an inversion channel) forms between S and D so that a current can pass. The thickness of the inversion layer, and also the charge carrier concentration, can be controlled by the gate potential. Actually, characteristics such as shown in Fig. 3-18 are obtained. While there still is a similarity to pentode curves, the rising segment is less steep than that of FETs.

If the P-type slice is suitably doped, a conduction channel already forms for UG equals 0. Then, it is necessary to make UG negative for ID cutoff. (This is the case in Fig. 3-18.) MOSFETs with gate and drain voltages of the same polarity are said to be **enrichment** types, while **depletion** types require opposite G and D polarities for ID cutoff. Besides our N-channel type, there are P-channel devices, too; they are just complementary. FETs being symmetrical by design, S and D can be inverted.

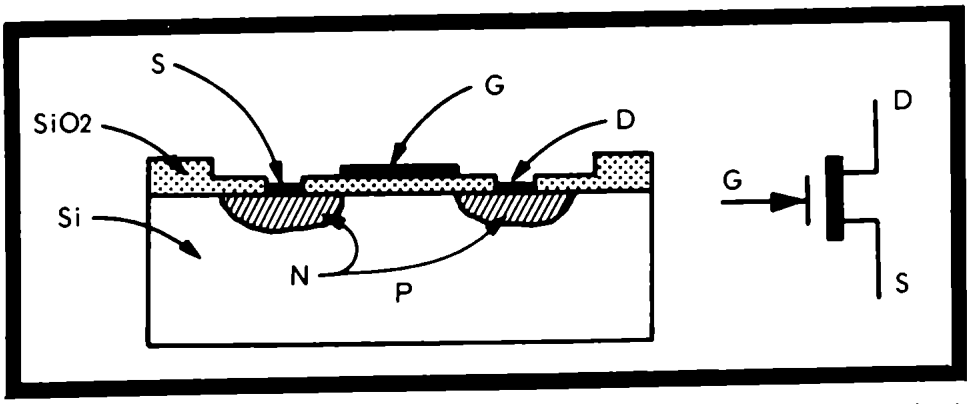


Fig. 3-17. Cross-sectional diagram and schematic symbol of a MOSFET.

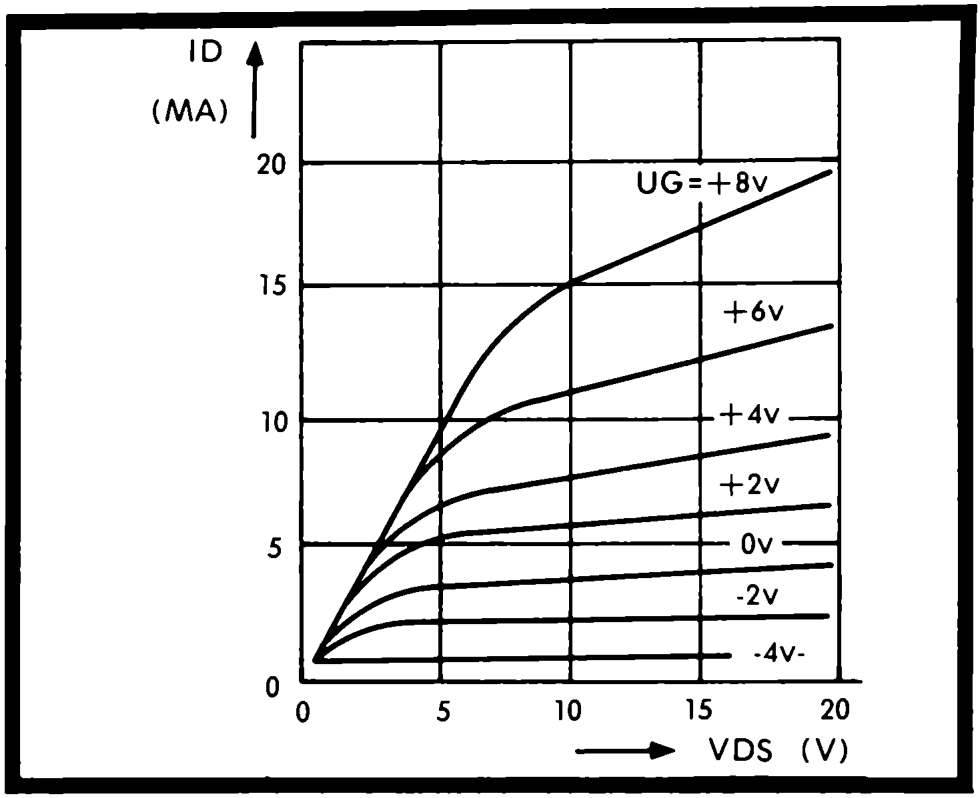


Fig. 3-18. MOSFET drain current (ID) drain-source voltage (VDS) characteristic.

MICROELECTRONICS

Thanks to the electron tube, semiconductor device transition, the volume of any given electronic equipment has shrunk in amazing proportions. The event of microelectronics is resulting in another equally significant reduction in bulkiness. As electronic gear goes increasingly sophisticated, its size has to be drastically squeezed, using new manufacturing processes to maintain the bulk of the whole within acceptable limits. This is true for the various Apollo spacecraft on-board computers, as well as for office desk-type calculators. **Monolithic integrated circuits** are thus produced by partially automatic processes, making possible price slashes that opened new outlets to the soon competitive ICs.

The MOSFET structure in Fig. 3-17 is made by IC-techniques, where a silicon single crystal slice, generally P-type, serves as the substratum, which is treated by chemical and physical processes such as etching, masking, photoetching, oxidation, reduction, diffusion and contact area vapor

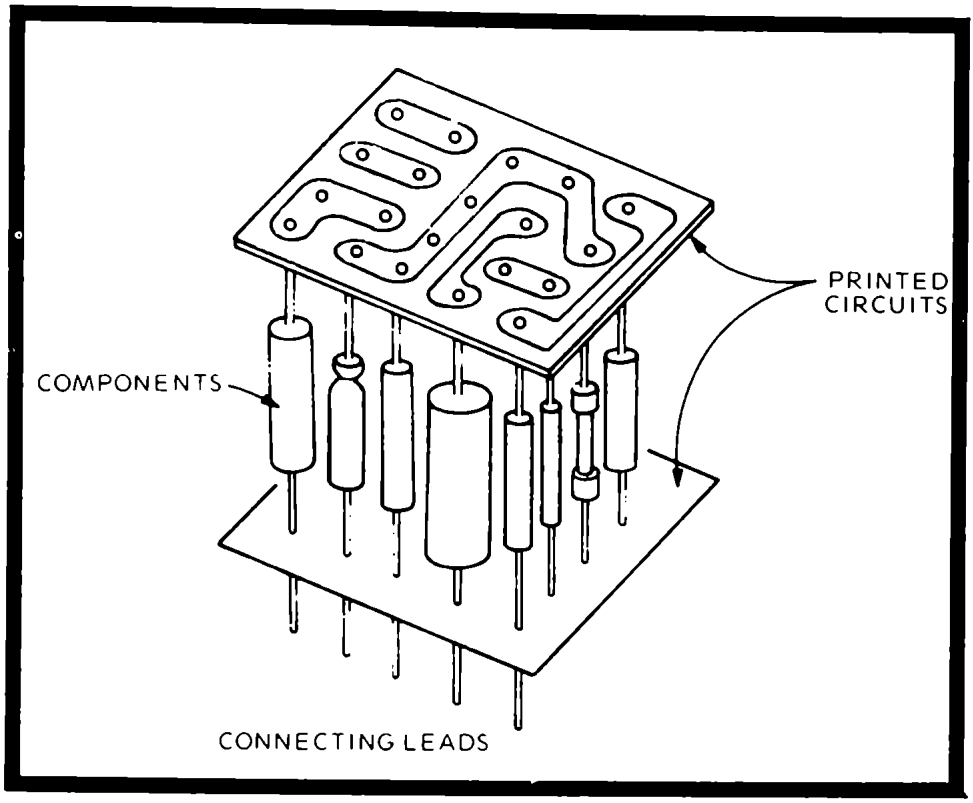


Fig. 3-19. The cordwood assembly technique features high component density.

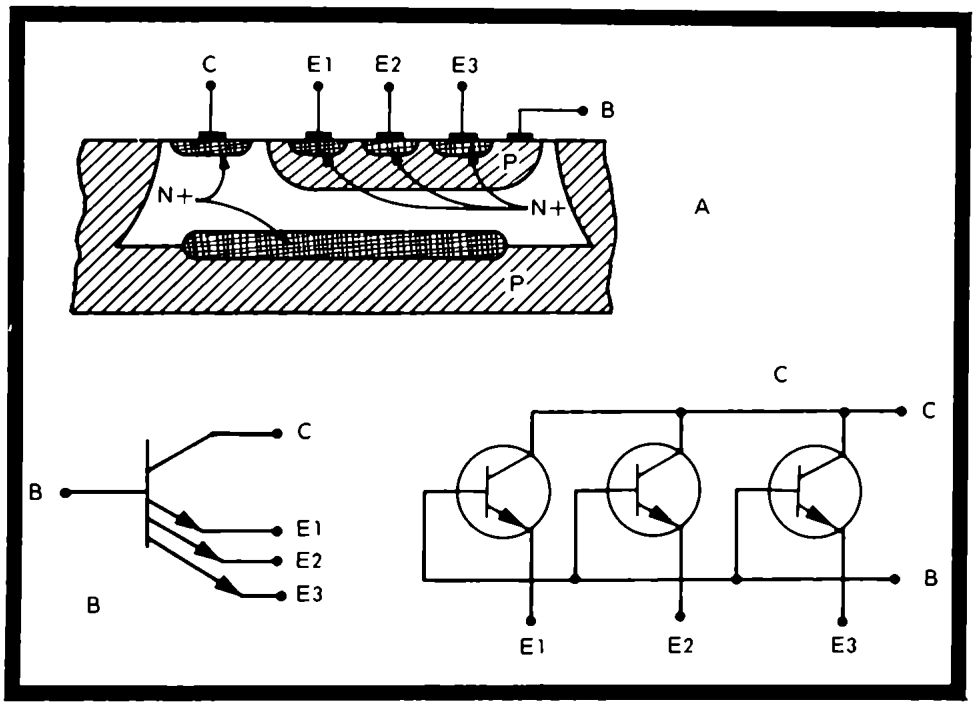


Fig. 3-20. Cross-sectional diagram, schematic symbol and equivalent circuit of a multi-emitter transistor.

deposition, all on a microscopic scale, for the slice contains a number of identical circuits. The finished slice is broken along fracture lines into individual circuit chips measuring from 0.02 x 0.02 to 0.08 x 0.08 inch. This is a very special technology, out of the scope of the present book. While integrated transistors are more or less similar to discrete devices, there are, however, special integrated components such as the NPN multi-emitter transistor in Fig. 3-20, equivalent to three transistors with the bases on one hand and the collectors on the other. This device is used in TTL logic circuits (Chapter 4).

Thus, it is possible to arrange on a single chip conventional NPN and PNP transistors as well as MOSFETs, resistors from, say, 10 ohms to 30K and capacitors of not more than about 300 pf. With such components, complete circuits for simple functions such as logic can be laid out—gates, flip-flops, scales-of-two, memory cells, etc., a **Digital IC**. The equipment manufacturer assembles and connects these circuits to build complete counter decades, code converters or other gear of a higher degree of complexity. But circuit manufacturers looking for more vertical business integration have started "integrating ICs" by arranging on a larger silicon slice the number of elementary circuits necessary to perform evolved functions such as a decade counter, formerly assembled by the equipment manufacturer. Thus, we had **medium scale integration** (from, say, 10 to 100 circuits per slice), and then came **large scale integration** (more than 100 circuits).

Such techniques, obviously, are economically justified only if large quantities of such MSI or LSI circuits are required, and this seriously limits their possibilities. An interesting solution may come from a technique called programmed connection control. The manufacturer prepares and stocks semi-finished slices containing a large number of unconnected identical transistors. According to requirements, a computer establishes the circuit diagram, tests the transistor elements on the chip, selects the best suited ones (the bad ones are left aside) and then traces the connections that are good only for a given diagram and a particular chip. This method allows for custom-made ICs and for automatic components and circuit testing. But, even though computers work with astonishing speed, computer time is very expensive, and someone has to pay for it.

While monolithic circuits are sophisticated semiconductor devices, **hybrid ICs** are made up of an insulating substratum with deposited circuit elements. FETs and passive components are made by **thin film** techniques (the layer thickness is measured by the number of atom diameters), using vapor deposition or cathodic sputtering. Other components and the

inner connections are deposited by the so-called **thick film** technique, using serigraphy (silk-screen process). Transistor chips are mounted on the substratum without their casing, thus reducing bulk as well as cost. The finished hybrid is then potted or otherwise cased for protection, allowing access to connecting pins or sockets.

Linear Integrated Circuits

According to their operating mode, integrated circuits are divided into **digital circuits** as mentioned above and are operated by the presence or absence of a signal corresponding to binary numbers 0 and 1 or the functions **on** and **off**, and **analog** or **linear circuits** supplied with varying-amplitude signals, used with more-or-less type controls. This latter category comprises all kinds of amplifiers: wideband and bandpass, operational amplifiers (Chapter 5) and voltage or current comparators, generally of the differential-amplifier type.

In designing a general-purpose linear IC, manufacturers strive to make it as universal as possible. Besides the input, output and power supply terminals, connecting points for feedback and inverse feedback loops and for bypass capacitors are, therefore, provided to allow for modification of the fundamental amplifier characteristics. Input, and sometimes the output, too, can be symmetrical or asymmetrical as desired, so that the output can be in phase or out of phase regarding the input. Such an amplifier, actually, is serviceable everywhere, and younger engineers tend to make a habit of designing electronic setups around such universal ICs rather than devising and laying out a specific-purpose amplifier. Will electronics soon become the art of discerningly interconnecting black boxes containing linear ICs?

Supply voltages	+12 -6	+6 -3
Input resistance	40K	67K
Output resistance	200 ohms	300 ohms
Open-loop voltage gain	3600	900
Output voltage swing	+5.3v	+2.7v
Dissipated power	80 mw	17 mw
Common-made rejection ratio	95 db	95 db

Table 3-1. Operating characteristics of the SGS-Fairchild uA702A IC.

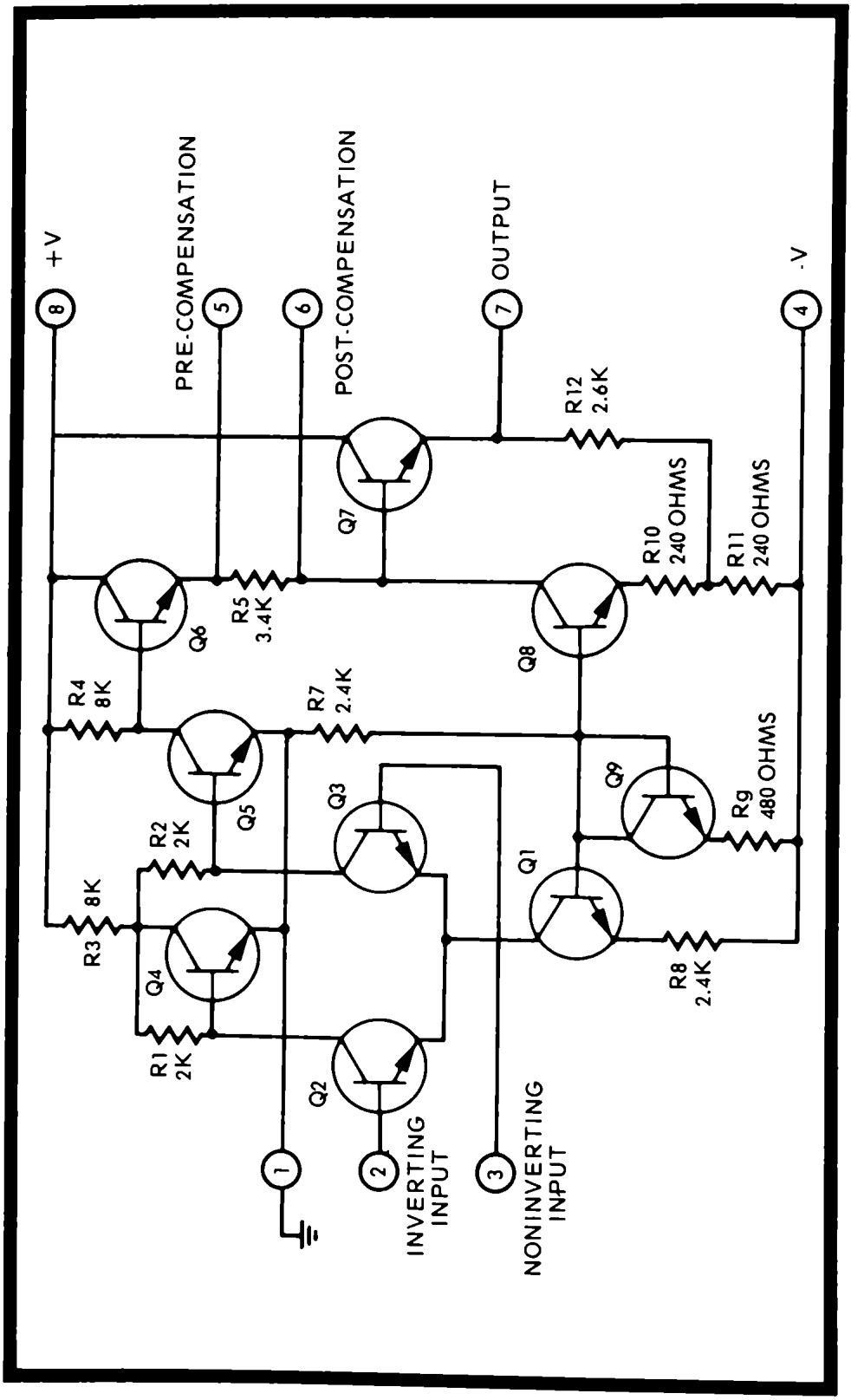
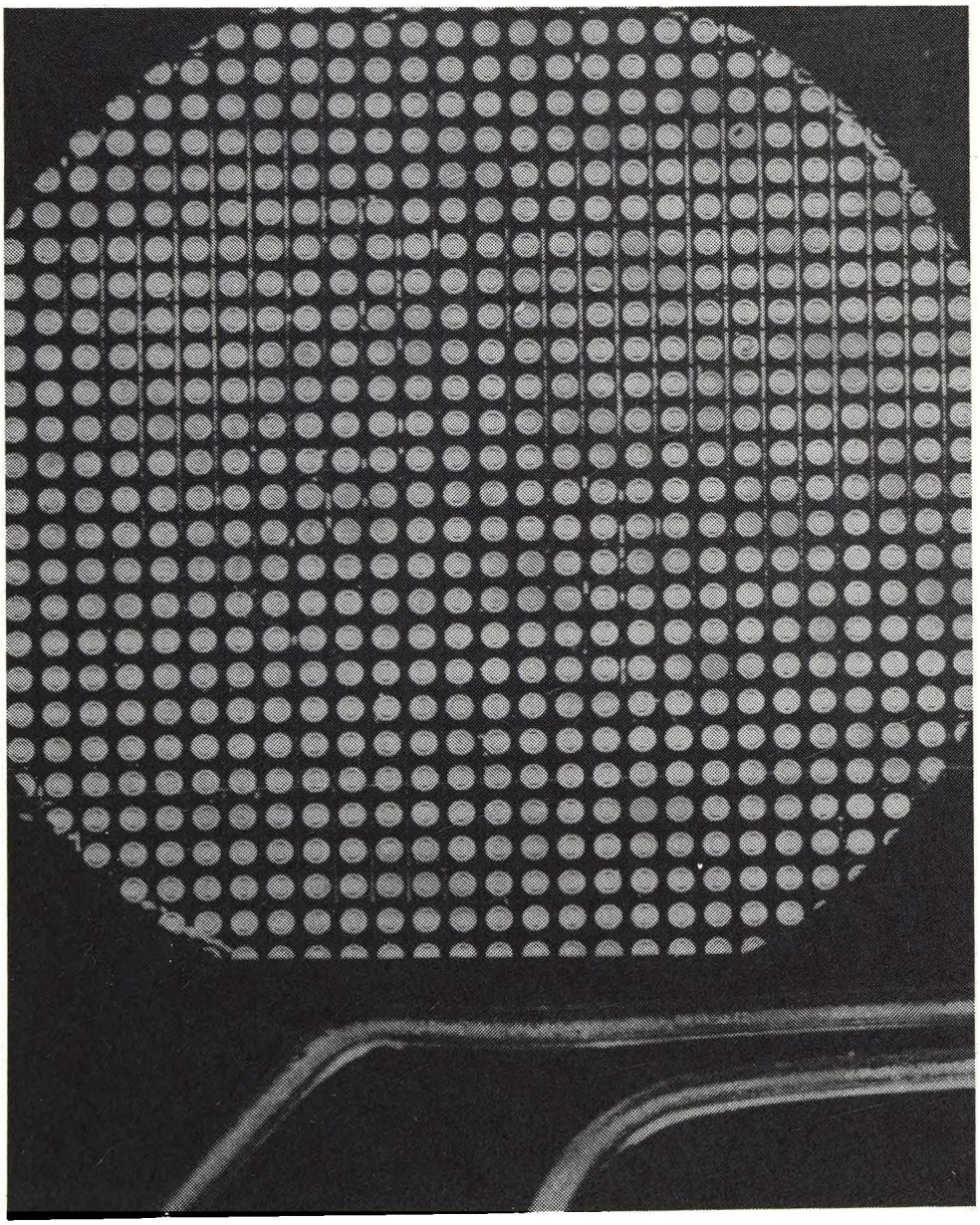


Fig. 3-21. Schematic of the SGS-Fairchild uA702A integrated circuit.

To cite an example, consider the circuit of the now conventional SGS-Fairchild integrated amplifier uA702A, comprising 9 NPN transistors and 12 resistors (Fig. 3-21). This circuit is composed of a high-gain differential input stage Q2Q3, fed by a constant-current source Q1Q9, a second symmetric input-asymmetric output stage Q4Q5, and an output amplifier, which includes a driver, Q6, fed by current source Q8 and a final CC-stage Q7. This amplifier, while only partially symmetric, is perfectly balanced, and with no input signal, the output connection is at ground potential. Table 3-1 lists some fundamental characteristics of this amplifier.



Two cm diameter planar slice containing some 600 chips.

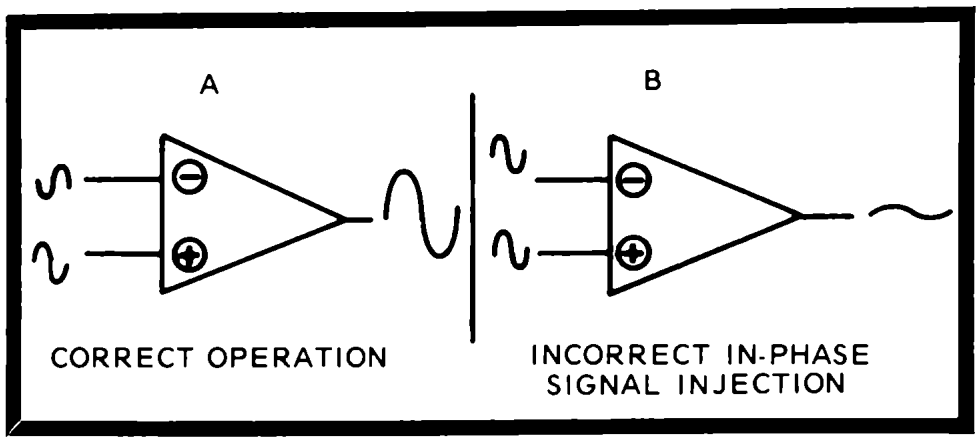


Fig. 3-22. As these diagrams indicate, a differential amplifier rejects common-mode signals.

The last characteristic in Table 3-1 illustrates an essential feature of symmetric-input amplifiers. Correctly operated, the inverting (-) and not inverting (+) inputs are driven by opposite signals to yield the specified gain (Fig. 3-22A). Undesirable signals, such as hum, will appear in phase on both inputs, or in a **common mode**, and by virtue of a high common-impedance (Q1Q9) differential amplifier, they are rejected by as much as 95 db and, therefore, do not appear at the output (Fig. 3-22B). This offers an outstanding facility for connecting high-impedance (necessarily symmetric) transducers free of hum and noise to the accompanying input amplifier.

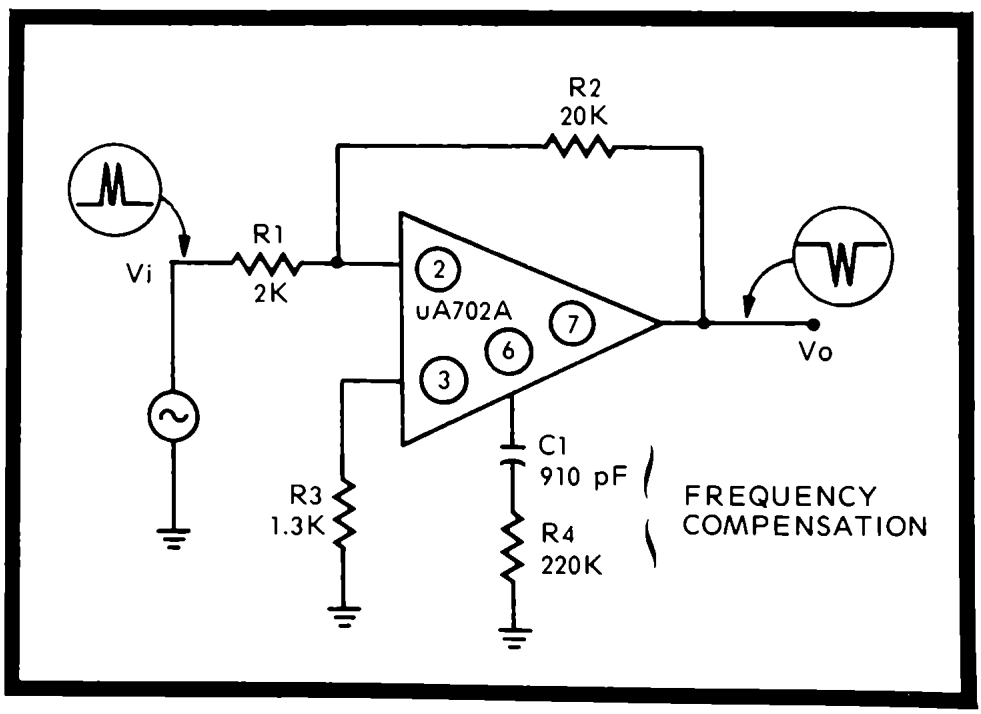


Fig. 3-23. Signal inverting amplifier IC connections.

To illustrate the amazing variety of uses for a typical linear circuit such as the uA702A, and also the multiple ways of adapting it, we will examine several application diagrams. The triangle is the accepted symbol of an amplifier. It is positioned like an arrow aimed at the output direction. The connecting point numbers refer to the corresponding circuit terminals in Fig. 3-21.

With the signal-polarity inverting amplifier shown in Fig. 3-23, input voltage Vi is applied to inverting input Terminal 2. The high open-loop gain G cited above is rarely used, principally because of the need to apply suitable inverse feedback to increase amplifier stability. Resistor R2, connected between Points 2 and 7, closes an aperiodic inverse feedback loop, reducing phase distortion and, thereby, instability due to phase rotation at roll-off frequencies; gain is reduced simultaneously to about 10. Resistor R1, series-connected with Input 2, increases the input resistance seen by the amplifier. R3 reduces the offset voltage, producing an output voltage with no input present. The series circuit, C1R4, inserted between Point 6 and ground, introduces frequency compensation by attenuating the higher frequencies, thus reducing bandwidth and noise, too. This method is said to be post-compensation; the overall inverse-feedback by R2,

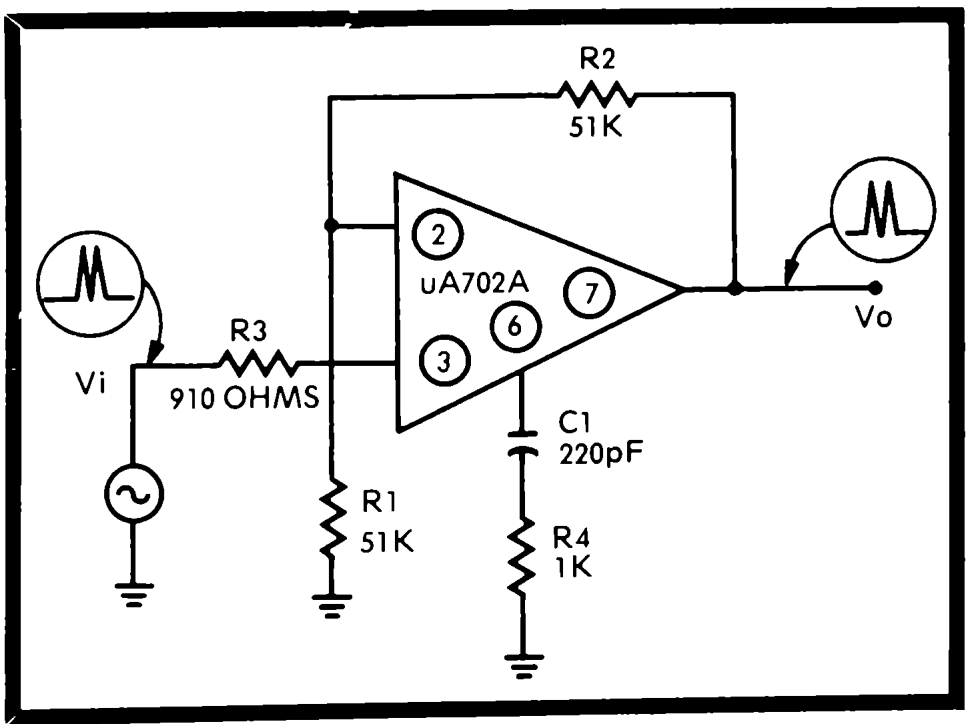


Fig. 3-24. In a true or noninverting amplifier, the input and output signals are in phase.

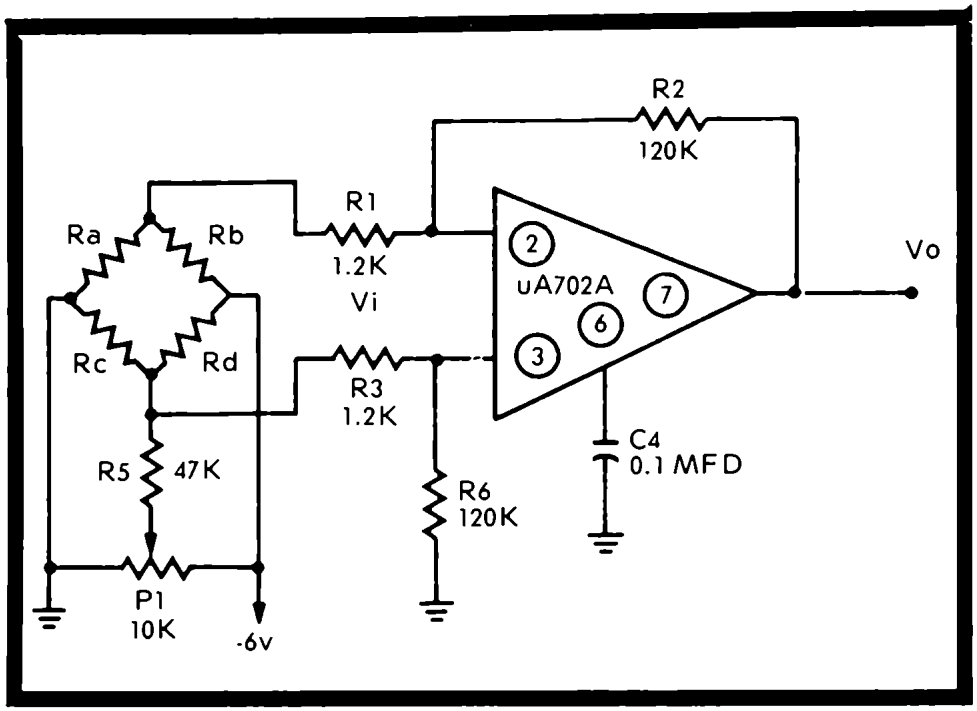


Fig. 3-25. Wiring for a differential input bridge null detector.

widening the bandwidth to again increase stability, is called **pre-compensation.**

The same circuit elements are found again in Fig. 3-24, but since Vi is applied here to Terminal 3, there is no inversion and Vo is in phase with Vi. The inverse-feedback loop, of course, remains tied to 2. The gain is mainly determined by the inverse-feedback resistors:

$$G = V_o/V_i = (R_1 + R_2)/R_1$$

R3 reduces the input offset voltage as well as its drift.

In Fig. 3-25, our do-all IC is used as a differential input null detector for a Wheatstone bridge; therefore, its inputs have been made resistively symmetric, choosing R1 equal to R3 and R2 equal to R6. For best results, the bridge arm resistances should be of the same order of magnitude, 120 ohms being an optimum value. Though the detector bridge diagonal is at a potential of about -3v, this voltage does not appear at the output because of the common-mode rejection feature and can be forgotten.

The AC amplifier in Fig. 3-26 offers a high input impedance, due to a feedback loop extending from 7 to 3 via R2, C2 and R3; this is called a **bootstrap** circuit. With the com-

ponent values as indicated, there is a gain of 10 and an input resistance of 2 meg, allowing for use with high-impedance transducers.

CARRIER-CURRENT AMPLIFIERS

Direct-coupled amplifiers, be they solid-state or tube-type, are fundamentally plagued by a **drift**, and drift is usually more pronounced with higher gain. Of course, there are measures to be taken to reduce drift, but they can be costly. So it may be more economic to design a carrier-current amplifier, although it is more sophisticated at first glance.

In a carrier-current amplifier, the slowly varying voltage that makes up the signal is chopped by means of a suitable chopping modulator (a vibrating-reed chopper, for instance). Thus, a wave made up of positive or negative pulses is obtained, and its envelope represents the original signal. Such a wave is easily amplified by a conventional AC amplifier, preferably tuned to the chopping frequency f to enhance the signal-to-noise ratio. Using an inverse-feedback loop, the gain of this amplifier can be maintained constant without any power supply stabilization.

Once amplified, the chopped signal voltage can be rectified; but if the signal comprises both polarities, the original sign is lost by mere detection. A phase detector or discriminator supplied with the same phase reference by the

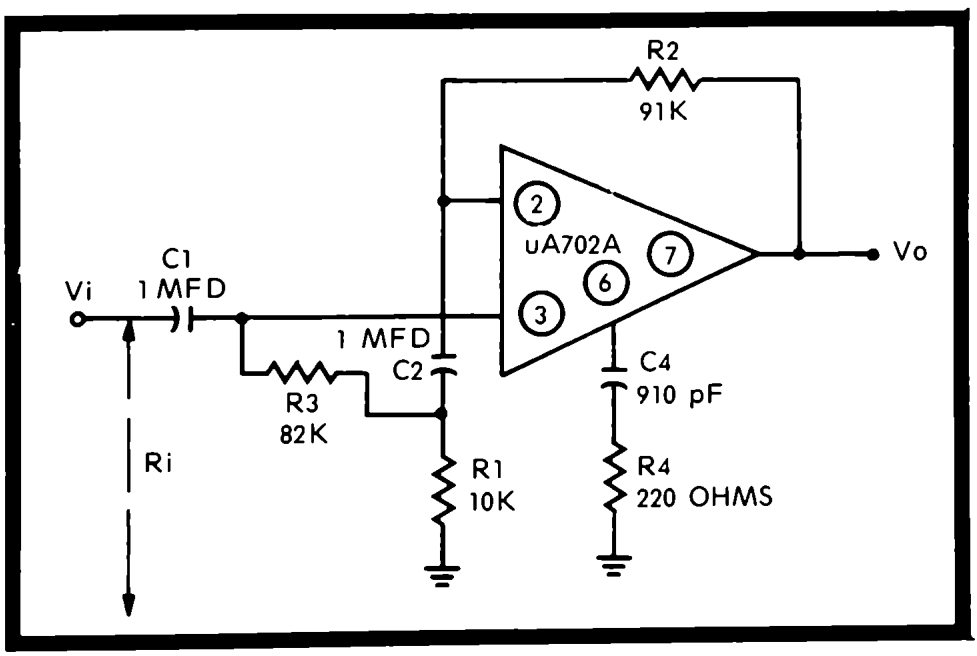


Fig. 3-26. AC amplifier with a bootstrap-type feedback connection.

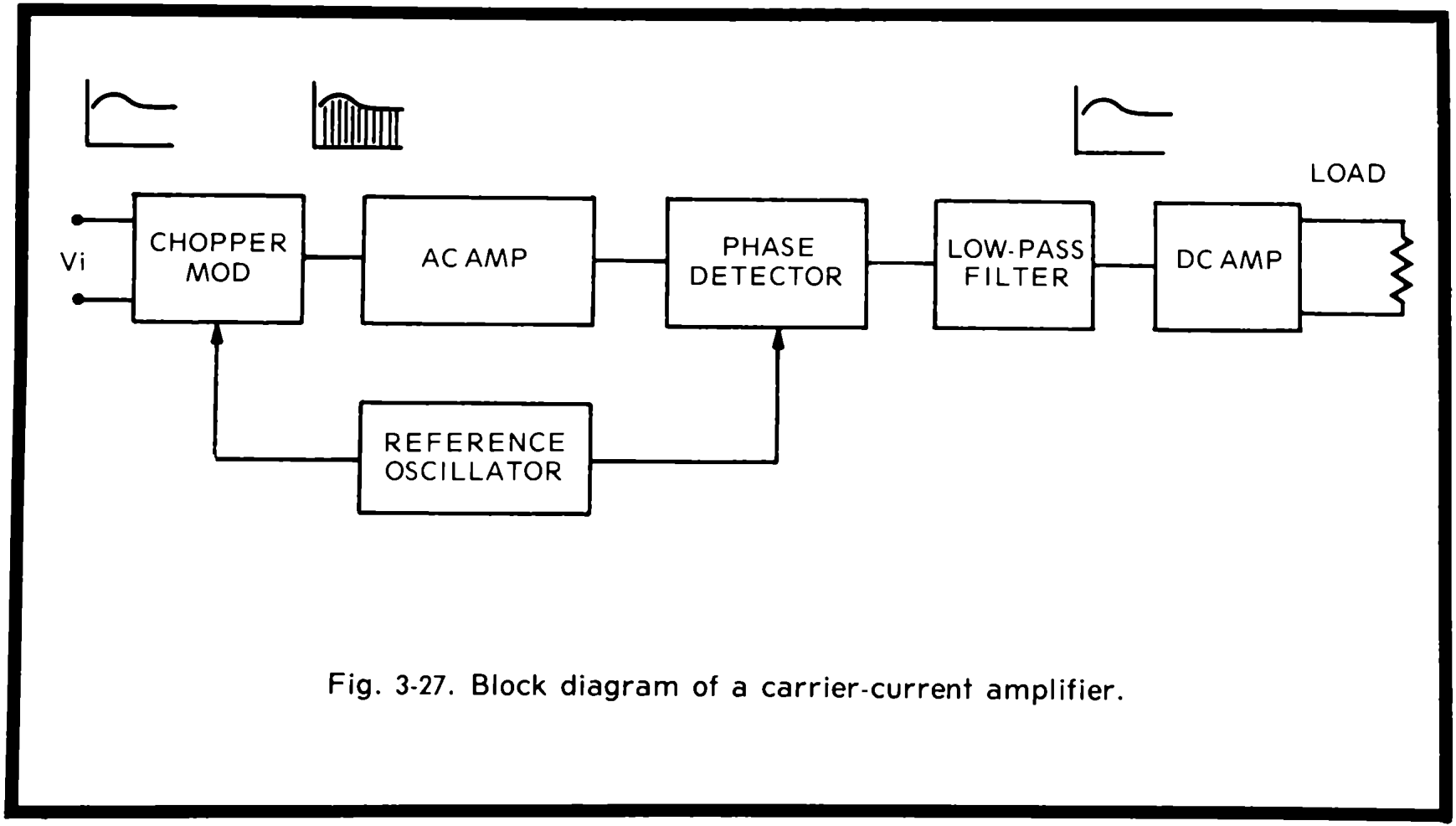


Fig. 3-27. Block diagram of a carrier-current amplifier.

chopper oscillator can restore the signal polarity, yielding the direction of transducer unbalance. A low-pass filter will eliminate residual modulation-frequency components. Now, a DC power amplifier stage can be used to drive the load (Fig. 3-27). Since the gain of this last stage is low, its drift can be negligible.

To outline the possible solutions, suppose we use resistive transducers such as strain gauges, potentiometer devices or thermistors, bridge-connected with a graphic recorder to record the variations of the parameter under control. Since the bridge error voltage is far too small to drive the sturdy recorder galvanometer, amplification is mandatory. As shown in Fig. 3-28, there is a choice between three solutions (all restoring the original signal polarity).

In Fig. 3-28A the bridge is DC-fed, and the error signal is applied to a DC amplifier. This solution seems quite easy, but drift is difficult to eliminate, especially if the gain is high. The bridge is still DC-fed in Fig. 3-28B, but here the error signal is chopped, amplified by an AC amplifier and detected with a phase reference by a discriminator, followed or not by a DC output stage. This circuit has no noticeable drift.

There is still another solution (Fig. 3-28C) where the bridge is fed an AC voltage of frequency f (from 60 to 10,000 Hz) generated by a reference oscillator. (Capacitors C' and C'' may be necessary to compensate for an unwanted reactive component.) The AC error voltage now is directly applied to the AC amplifier input, and then comes the phase discriminator with its reference phase injection and a possible output stage as before. Chopper and DC bridge supply are not necessary in this system, and drift is negligible, too.

CHOPPER MODULATORS

There are many types of chopper modulators suitable for control applications. Probably the oldest one is the electromagnetic vibrating reed chopper, which is a biased relay, sufficiently quick to be operated at a supply frequency between 50 and 400 Hz (Fig. 3-29). The vibrating armature supports one or more sets of double-throw contacts. Since it is magnetically biased by a small permanent magnet, contact always is made on the same side for positive half waves and on the opposite side for negative half waves; this yields the phase reference.

The circuit diagram illustrates the use of the reed chopper with a center-tapped transformer. The voltage to be measured is successively applied with reversed polarity to both halves of the transformer, and the secondary develops an ap-

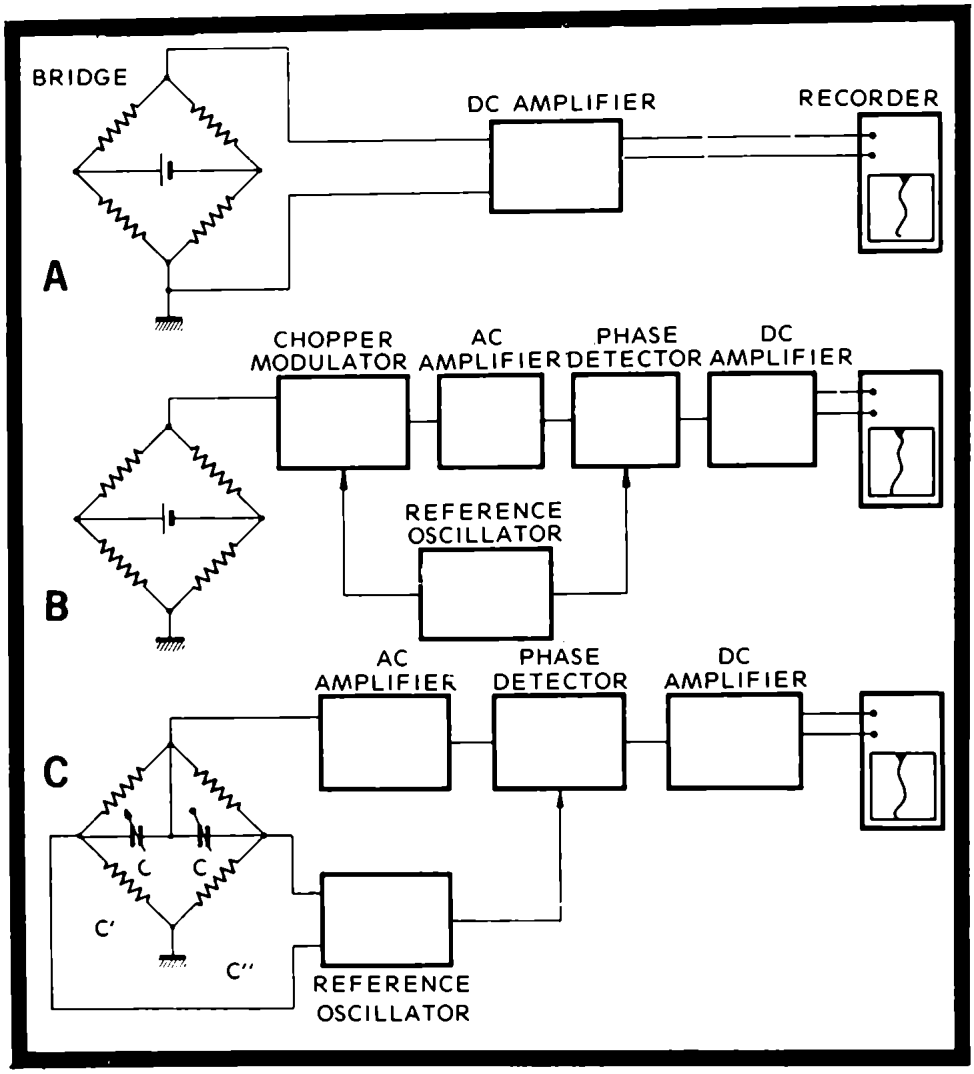


Fig. 3-28. Each of the three amplifiers shown will restore the polarity of the original signal after AC amplification.

proximately square wave with an amplitude proportional to the signal amplitude.

Metallic contacts look fine because of their low and well defined contact resistance; however, they are noisy and subject to wear. So contactless choppers are preferred. The capacitive chopper in Fig. 3-30 is electromagnetic, too; its mechanism is similar to the moving coil "motor" of a conventional speaker, driving a mobile insulated armature (A), which is part of a parallel-plate capacitor (C), rather then a membrane. A stationary armature (B) facing A (well insulated, too) forms the input electrode. As the charge (Q) of a capacitor equals its capacitance (C) times the applied voltage (V), a variation of C, with Q remaining constant, results in a variation of V, and across series resistor R develops an ap-

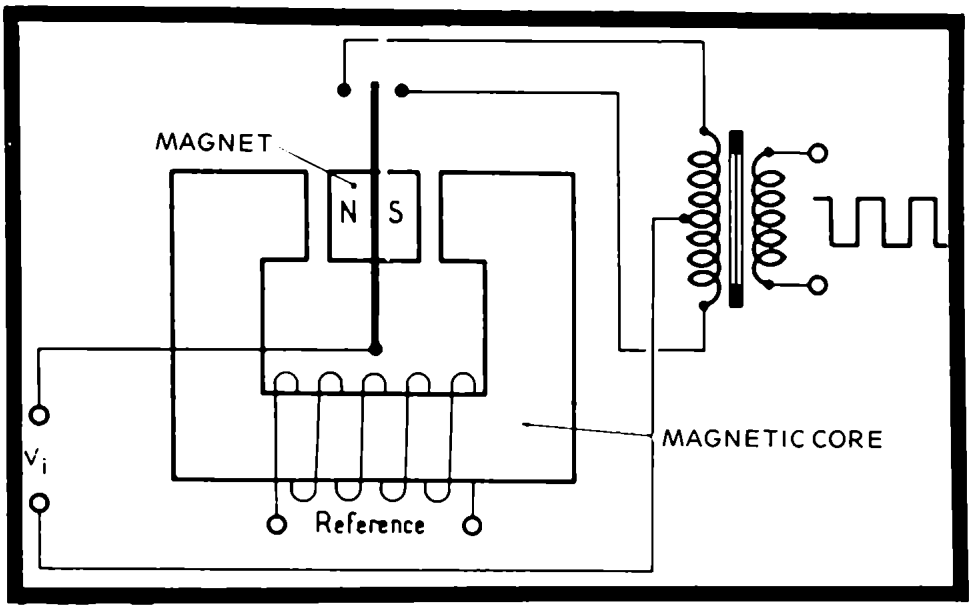


Fig. 3-29. Diagram illustrating the operation of a vibrating chopper modulator.

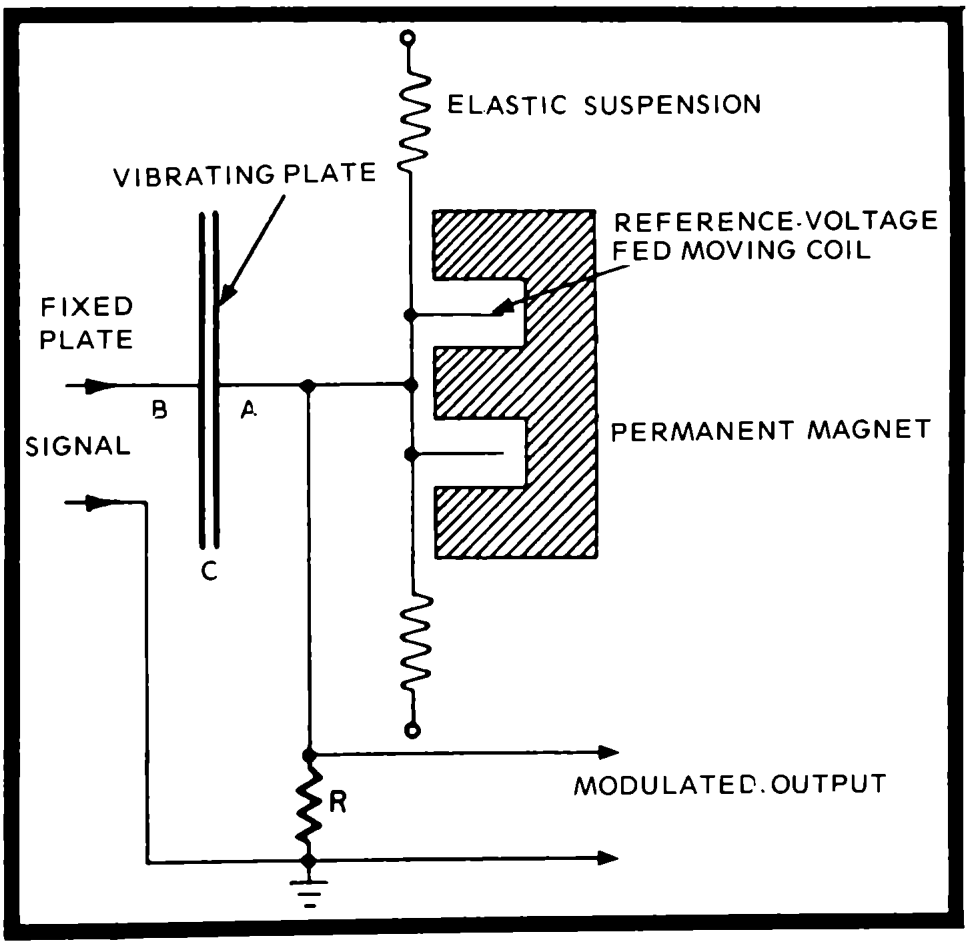


Fig. 3-30. A capacitive modulator uses a "motor" similar to that found in a dynamic speaker.

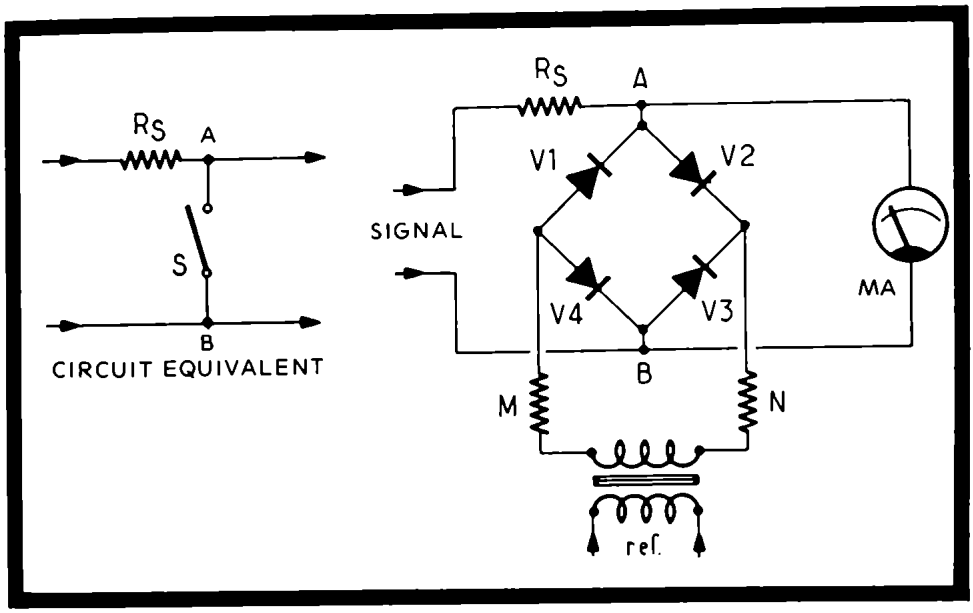


Fig. 3-31. Bridge modulator and its equivalent switch circuit.

proximate sine wave with an amplitude proportional to the signal. This capacitive modulator features an extremely high impedance; therefore, it is suited for special applications such as pH measurements.

Semiconductor technology has brought forth static solid-state chopper modulators that are extremely compact and rugged (and silent, too). Since there are no moving parts such choppers should never wear out. The **bridge modulator** (Fig. 3-31) is composed of four matched diodes connected two by two, just as in some instrument rectifiers. The DC (galvanometer) terminals are supplied the reference voltage, and the AC diagonal is used as a switch, shorting Points A and B behind series resistor Rs. During half cycles where M is positive in regard to N, the four diodes are conducting, and the resistance of bridge diagonal AB is low in comparison to Rs, just as if the switch of the equivalent circuit were closed. During half waves of opposite polarity, the diodes are nonconducting (for silicon diodes, resistance then exceeds 10 meg), and the signal is transmitted with virtually no attenuation.

Replacing diodes V3 and V4 by matched resistors results in the half-bridge modulator in Fig. 3-32. Potentiometer P is necessary to balance left- and right-hand resistance arms of the bridge to avoid having the reference voltage reach the output terminals during diode conduction.

Transistors make good base-current controlled chopper modulators. With no base current, the collector current of the transistor in Fig. 3-33 is zero and the switch is open. Applying a

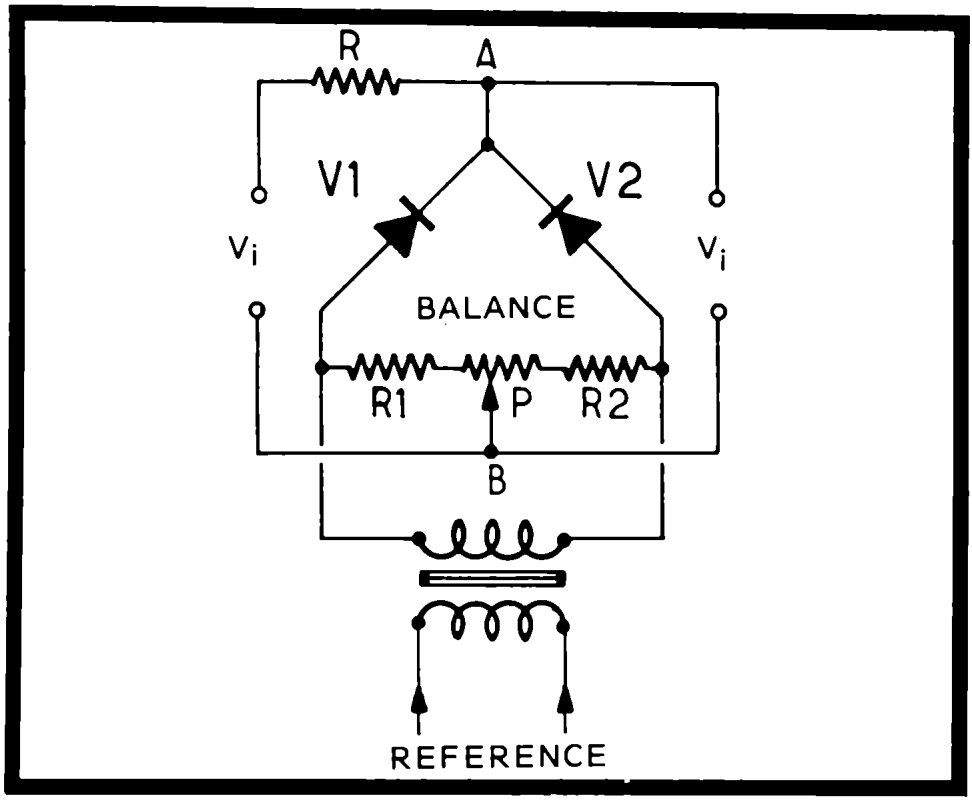


Fig. 3-32. Half-bridge modulator or phase detector.

negative base voltage (V) results in collector current saturation, and the transistor behaves as a low resistance and short-circuits the signal.

The phase detector, fed by TR1, is a bridge modulator such as shown in Fig. 3-31, but is is a full-wave type, using two identical windings on reference voltage oscillator transformer TR2 and two identical though differently connected rectifier bridges. The rectified output current traverses the load (represented here by a 20 ma 2K galvanometer), and a 600-

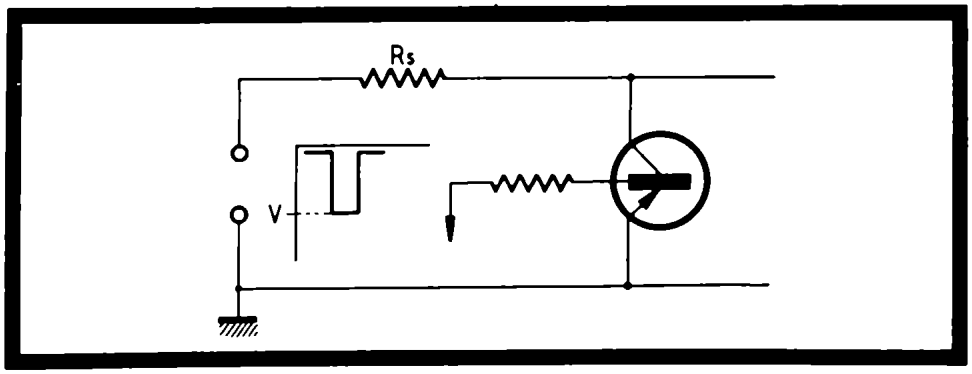


Fig. 3-33. Switching transistor circuit designed for use as a modulator.

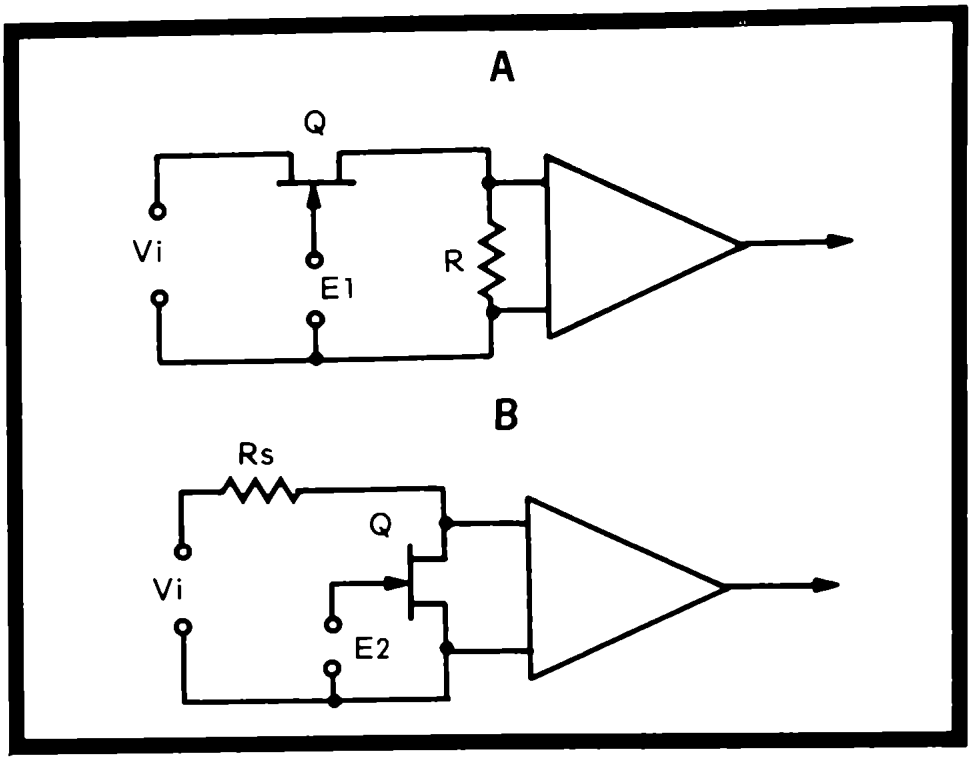


Fig. 3-34. Series (A) and parallel-connected FET chopper circuits.

ohm resistor, Rc, series-connected with the input signal. Rc develops an inverse-feedback voltage only about 0.1 percent smaller than the signal voltage. This makes for the extremely low current drain of the amplifier: only 3 microamperes for a 12v input, that means an input resistance of 4 meg, M a good example of a clever application of inverse feedback in view of an important input resistance increase. The value of Rc is critical; therefore, a wirewound precision resistor is used. The 500-Hz reference oscillator is composed of push-pull connected transistors T7 and T8.

For the chopper function, an FET is much better suited than a (bipolar) transistor because its drain-source resistance can be made to vary between a minimum of 50 ohms and 10K (according to type) and a virtually infinite value by applying a couple of volts to its gate. The FET modulator can be series-connected (Fig. 3-34A) or parallel-connected (Fig. 3-34B). In the series mode, input voltage Vi develops across resistor R (input resistance of follow-up amplifier) when the FET is conducting; when it is not, there is no such voltage. In the parallel circuit, FET Q and series resistor Rs form a voltage divider in which parallel component Q periodically forms a short. In both cases, a chopped wave appears across the circuit output if Q is fed a rectangular wave.

Outstanding efficiency is obtained with the series-parallel chopper modulator in Fig. 3-35. FETs Q1 and Q2 are alternately made conductive by out-of-phase control voltages e1 and e2, and the amplifier input voltage continually switches between Vi and 0.

The presence of a spurious gate-channel capacitance in FET modulators results in an undesirable amplifier input transient because the steep front of the switching wave overshoots. Therefore, some type of neutralization is needed to compensate for the resulting spurious component, perhaps an out-of-phase control voltage injected into the circuit through an adjustable capacitor. This compensation, however, is not perfect, for the spurious capacitance to be neutralized is that of a reverse-biased diode, which varies with applied voltage. MOSFETs feature a true capacitance that will allow compensation. Therefore, they are preferred to FETs in chopper modulator design.

Opto-electronic chopper modulators are described in Chapter 2. The FET equivalent of the parallel-type chopper in Fig. 2-59 appears in Fig. 3-34B, and the series-parallel opto-electronic device FET counterpart is shown in Fig. 3-35. Opposite to FETs and even more to MOSFETs, such opto-electronic couplers can resist practically unlimited potential differences between control and output circuits. The former devices, using slow response glow- and incandescent-wire lamps and photoconductors, are being superseded by all solid-state devices composed of a light-emitting gallium-arsenide (GaAs) junction diode or some other type, and a phototransistor or other light-sensitive function device, enclosed in a compact plastic housing made opaque to shield against outer illumination (Fig. 3-36). The light-emitting and receiving junctions may face eacn other, or they may be coupled by a light-conducting medium such as optical fibers. Electrostatic coupling may be eliminated by interposing a wire screen.

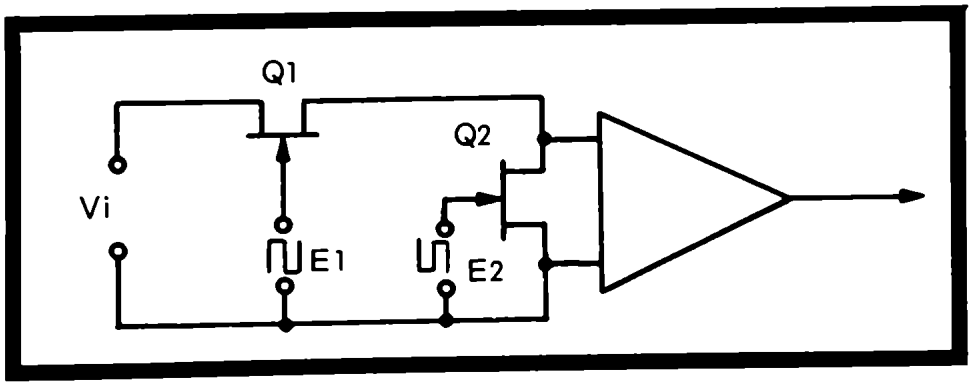


Fig. 3-35. Highly efficient series-parallel FET chopper modulator circuit.

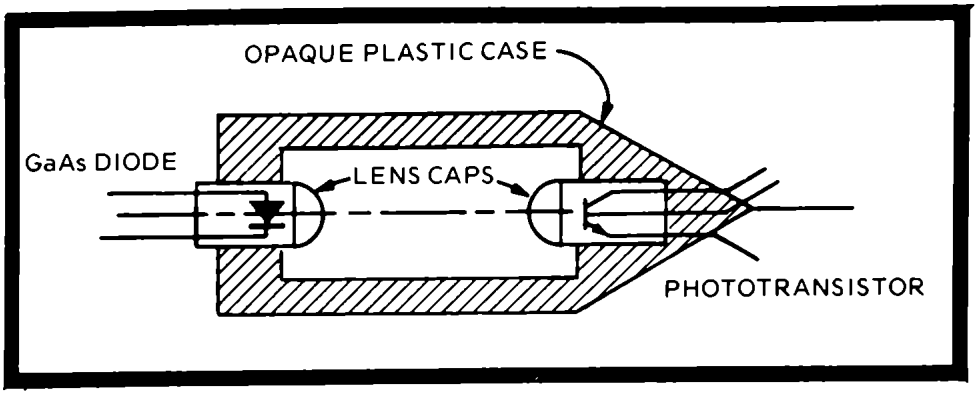


Fig. 3-36. Diagram of an opto-electronic chopper.

The **Hall effect modulator** (Fig. 3-37) also features very good insulation between signal and reference circuits. The DC or slowly varying signal voltage Vs applied to one pair of transducer electrodes is modulated by the reference current induced magnetic field set up in the air gap of an electromagnet, and the modulated output voltage, Vo, develops across the other electrode pair. This modulator is a low input and output impedance device (one to 10 ohms) with a noise level of less than 0.1 microvolt; it is serviceable from very low temperatures up to about 210 degrees F.

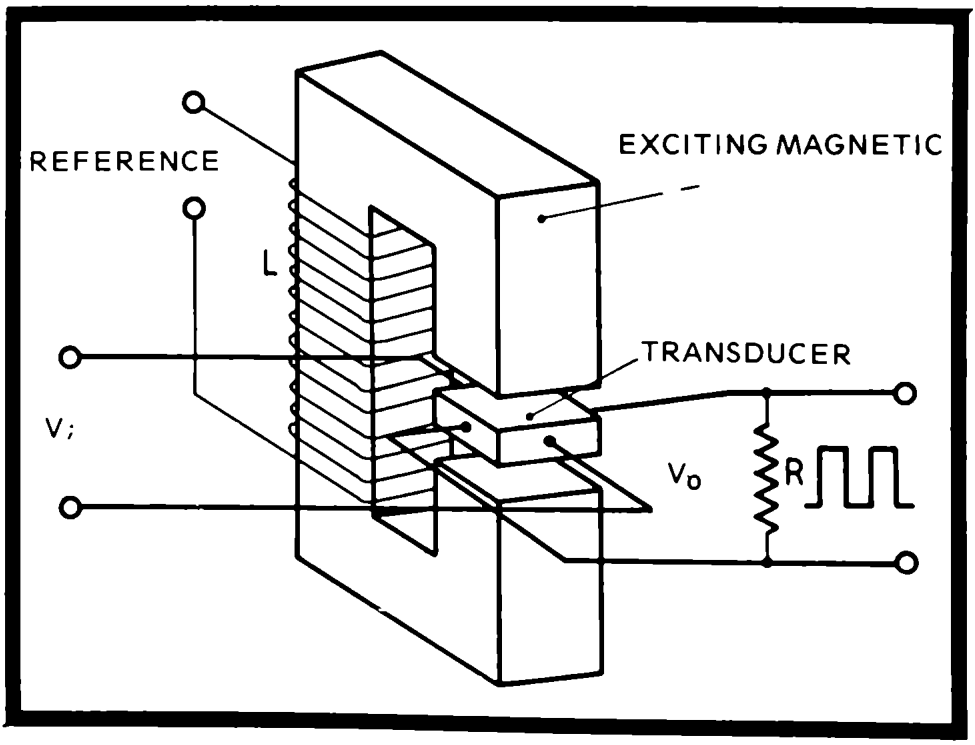


Fig. 3-37. Diagram of a Hall effect device used as a modulator.

SWITCHING BEHAVIOR OF THE TRANSISTOR

In its closed state, a good switch offers a negligible voltage drop Vk so that voltage V across load RL differs the least possible from supply voltage E (Fig. 3-38A). In its open state, the circuit current (I) has to be zero for V to be zero, too (Fig. 3-38B). In electronic controls, transistors and other semiconductor devices are used whenever possible, instead of relays, because static switches do not wear out and, therefore, require very little maintenance; and then, static devices may accommodate extremely high switching speeds such as required by computer techniques, while moving-contact devices have a limited operating speed by nature.

But while moving-contact devices have perfect switching characteristics, transistors are not such good switches. In the heavily conducting state (switch closed), there is a residual voltage Vk of some tenth of a volt, and in the cutoff state (open switch), there is a low though not negligible residual current. Furthermore, this current is temperature-dependent, especially in germanium transistors where the residual collector current increases about twofold for a temperature increase of 10 degrees C, or 18 degrees F. Silicon transistors and integrated circuits, however, offer low residual currents and reduced thermal variations thereof.

To understand the switching behavior of a transistor, consider its Ic equal to the f (Vce) characteristic shown in Fig. 3-39, where MQ represents the load line for definite operating conditions. Points M and Q define the "ideal" switch. At M, Vce (or voltage drop Vk) is zero, and at Q, Ic is 0. The actual operating points N (conduction) and P (nonconduction) have to come as close as possible to M and Q, respectively, for acceptable switch characteristics.

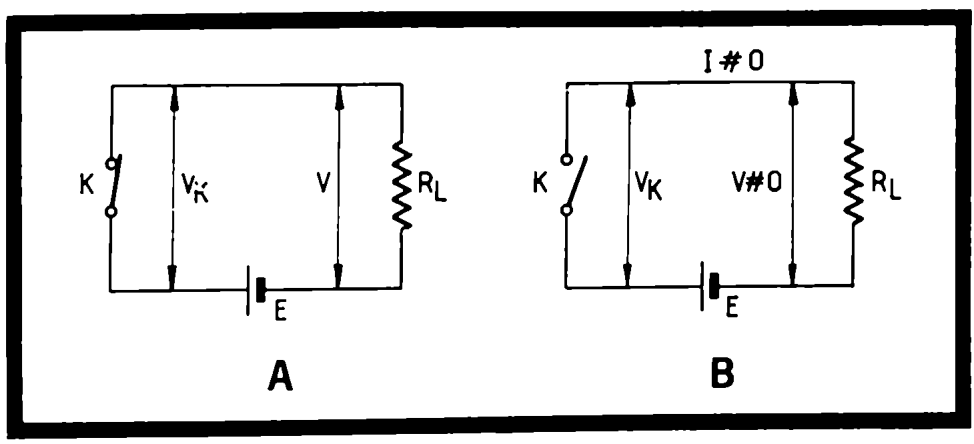


Fig. 3-38. With any switching device, there are two possible states.

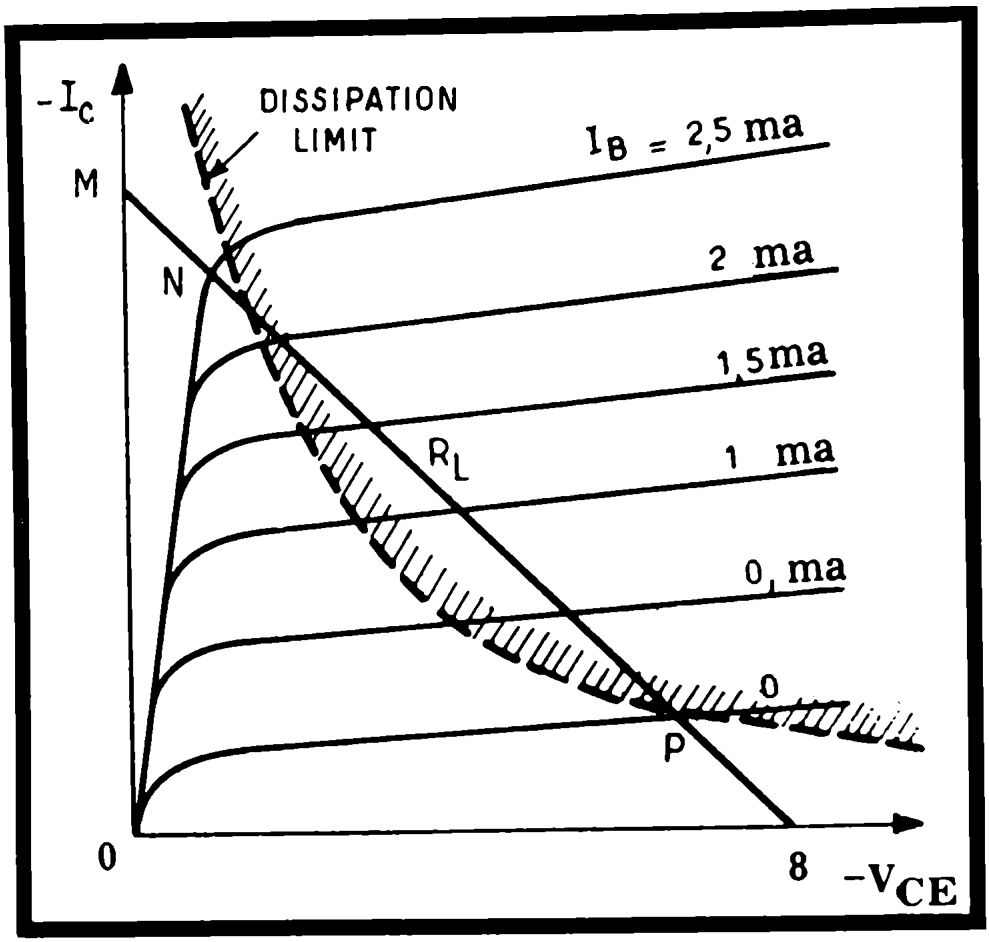


Fig. 3-39. The diagonal load line shows the best switching points of a transistor. The dashed line defines the dissipation limit.

Another important consideration is the rated power dissipation limit specified by the transistor manufacturer, shown traced on the characteristic. To the right of the dotted line, the power dissipated within the device would be excessive, leading to exaggerated heating and eventually destruction. Notice that N and P are situated within the permitted region where the transistor can operate indefinitely; but passing from N to P or from P to N, the operating point crosses the interdicted region inside the parabola. This is harmless if done quickly enough to avoid heating effects. Therefore, the transistor has to be fed a control signal of sufficient amplitude to sweep base current Ib at least from 2.5 ma to 0; furthermore, a steep-slope switching wave (preferably a square wave) must be used to achieve a rapid crossing of the dangerous region.

Transistors generally are more stressed during switching than in their steady state. Where large capacitors are used in

the circuit, potentially a dangerous current transient appears, and when a high inductance is disconnected, it may produce a voltage spike leading to junction breakdown. Vacuum tubes are less vulnerable to damage from such transients.

Before a slowly-varying control voltage can be applied to a switch circuit, it must be transformed into a steep-front wave for well-defined switching conditions. This is done by the Schmitt trigger, T1 and T2 in Fig. 3-40 (discussed later in this chapter) that yield a square-wave output for any input waveform of at least 4.5v of amplitude. This step voltage is transmitted to output transistor T3, by zener diode D, which is conductive only for inverse voltages overcoming some voltage threshold, the so-called zener voltage. This signal is transmitted to T3 only when the zener is in the conducting state; thus, there is no need to care for residual current in T2 when it is nonconducting.

For on-off control of relays and other power devices, thyratrons and thyristors may be more convenient to use, especially for all but very low power applications.

IMPEDANCE CONVERTERS

Some transducers, such as piezoelectric ones, have a high to very high impedance. Electrically, this means they behave like a generator (G) connected to the device terminals through a series resistor (R) of high value (several megohms). Con-

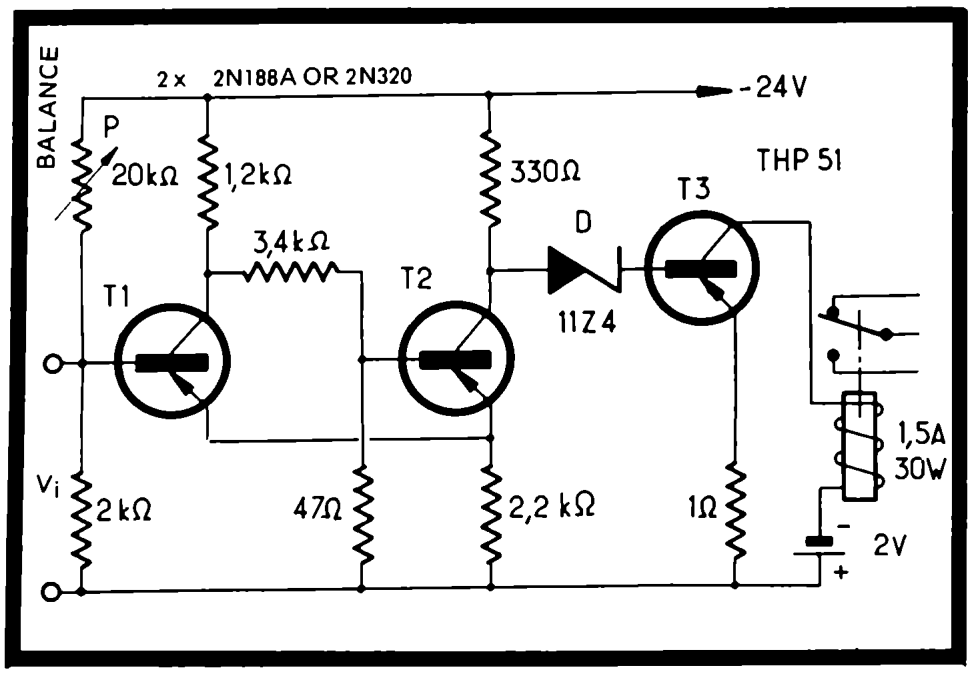


Fig. 3-40. A Schmitt trigger provides clean relay switching.

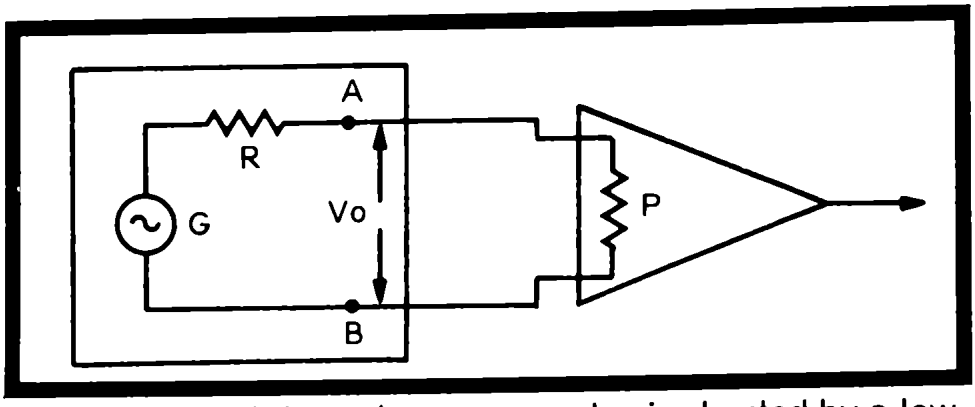


Fig. 3-41. A high-impedance generator is shorted by a low-impedance load.

necting a transistor amplifier of comparatively low input resistance r across such a transducer (Fig. 3-41) breaks down the output voltage due to r. Therefore, an impedance converter stage is needed between the transducer and amplifier input to sample Vo, with an input resistance Ri much greater than R, and deliver this signal with about unity gain across an output resistance Ro of sufficiently low value without distorting it.

An impedance converter using a triode (Fig. 3-42A) or a transistor Fig. 3-42B would suffice. The "cathode follower" circuit (Fig. 3-42A) features the high input impedance characteristic of vacuum tubes. Its output impedance practically equals 1 over S; S is the mutual conductance of the tube; with current tube types, output impedances range between 100 and 500 ohms. The voltage "gain" of the stage is 0.7 to 0.85, but there actually are definite current and power gains.

Regarding the usual CE configuration, the CC-connected transistor circuit in Fig. 3-42B increases the input impedance from, say, 1.3 to 350K for a given transistor type, while lowering its output impedance from 50,000 to 500 ohms. The voltage gain of 270 will fall to a little less than unity, and the power gain shrinks from 40 to 15 db; only the current gain remains about the same (between 35 and 36).

In the semiconductor domain, however, the FET is the ideal device for an impedance converter. The very simple fundamental circuit is shown in Fig. 3-43. Since the resistive input impedance component is extremely high, the capacitive component is the more noticeable; but with the stage gain slightly less than unity, the static input capacitance determined by design, is much larger than the dynamic input capacitance characteristic of the stage input and, thus, practically can be neglected (only 2 to 3 pf). This at least holds

for the common-drain connection used here; in the common-source configuration, this dynamic capacitance, on the contrary, is a multiple of the static capacitance and, therefore can be frustrating. The output impedance (Zo) of this "source follower" equals 1 over S, just as in the case of the cathode follower (neglecting source resistor R). It may be necessary to bias gate G positively or to tie R to a negative potential in order to allow for a sufficient input voltage swing.

FEEDBACK OSCILLATORS AND WAVE GENERATORS

The input and output signals of the tube or CE-connected transistor amplifier in Fig. 3-44 are out of phase; feeding back

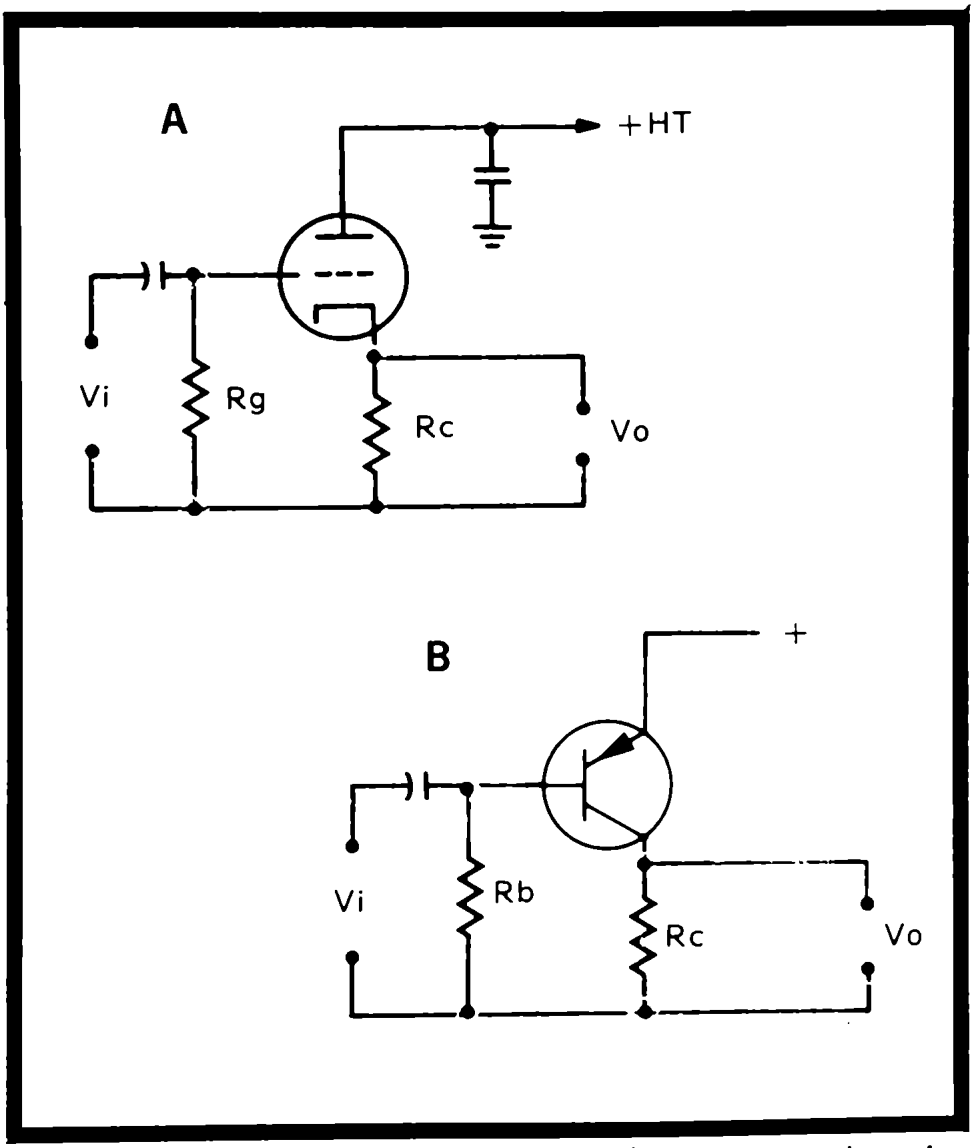


Fig. 3-42. Tube and transistor impedance converter circuits.

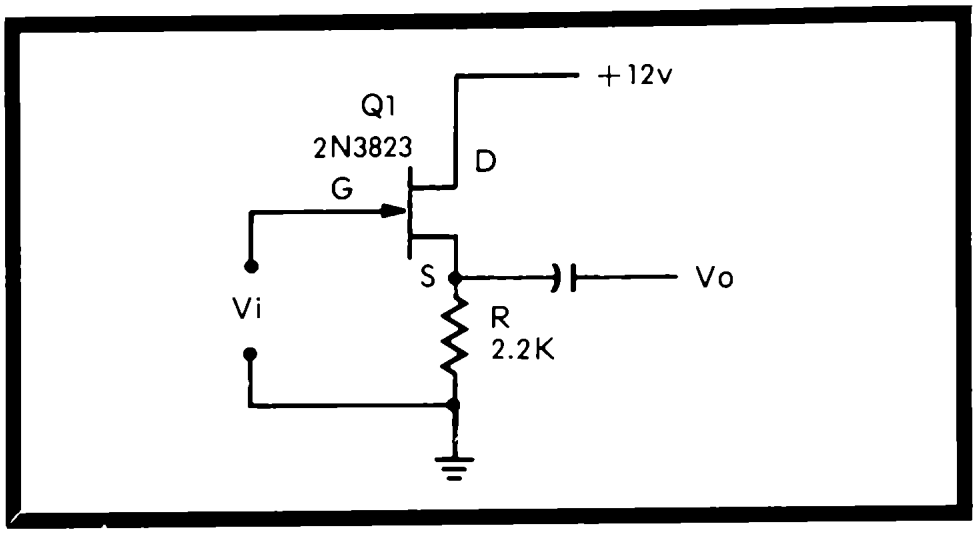


Fig. 3-43. FET impedance converter circuit. using the common-source configuration.

to the input part of the output, thus, would result in an inverse feedback loop and reduce the amplitude of a possible signal. However, inserting a network that introduces a 180-degree phase shift in the loop would add up to a complete phase shift of 360 degrees or 0. Now we have positive, signal-reinforcing feedback, and, supposing a loop gain higher than circuit losses, any signal will result in sustained oscillations. Oscillation amplitude will increase until something brings it to

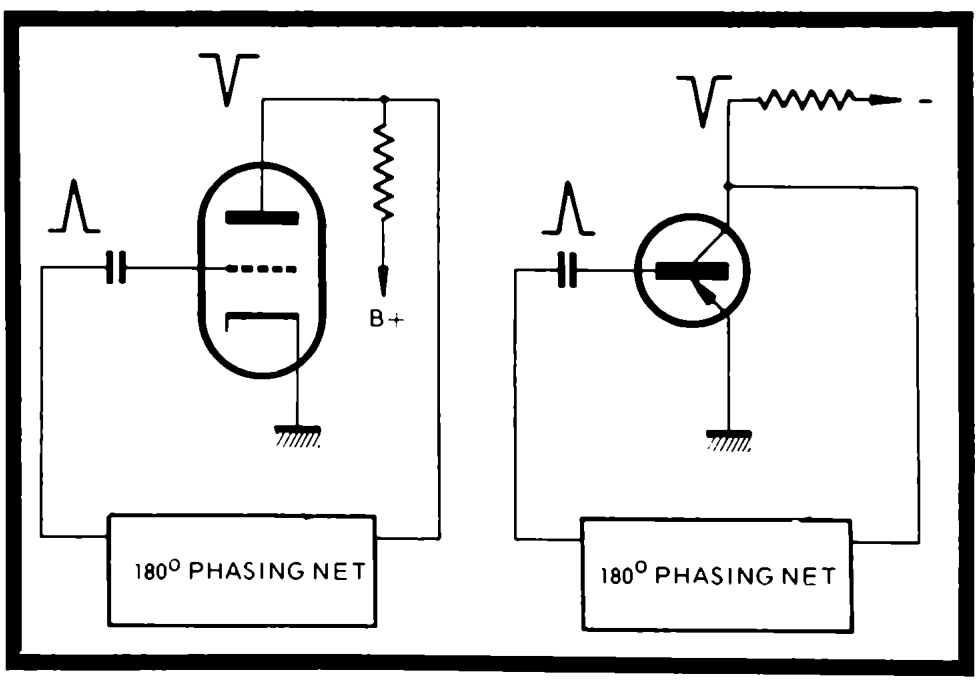


Fig. 3-44. A one-stage amplifier needs a phase-reversing network to oscillate.

a halt. The practical limiting factor is tube or transistor characteristic curvature, leading to an approximately square waveform.

In a 2-stage amplifier (Fig. 3-45), the input of V1 and the output of V2 are already in phase, and a loop containing a zero phase-shift network (a resistor for instance, or even a straight wire) is all that is necessary to start oscillations (adequate gain provided). With two tubes or transistors, there will be too much gain, and an inverse-feedback loop may be necessary to keep it under control. Insufficient gain would result in an unstable (ringing) amplifier.

Among all the waveforms, the sine wave is unique by the fact that it is composed of only one frequency component—the oscillating frequency f. (There are approximate sine waves obtained by distorting a "pure" sine wave; besides the fundamental frequency f, they contain multiple components such as 2f, 3f, 5f called harmonics.) To obtain a sine-wave oscillator, a selective network featuring a frequency-variable phase-shift is used and adjusted to get a 0 degree overall phase shift (amplifier and network) at the operating frequency f only. To avoid spurious oscillations, the loop gain is reduced as much as possible while sustaining oscillations.

To obtain a stable and well-defined oscillation frequency, it is necessary to use a selective device with a rapid crossover at the zero phase-shift line corresponding to f or, putting it otherwise, a circuit with a high Q factor.

The principal RC networks, serviceable for filter and oscillator circuits, are presented in Fig. 3-46. There are the

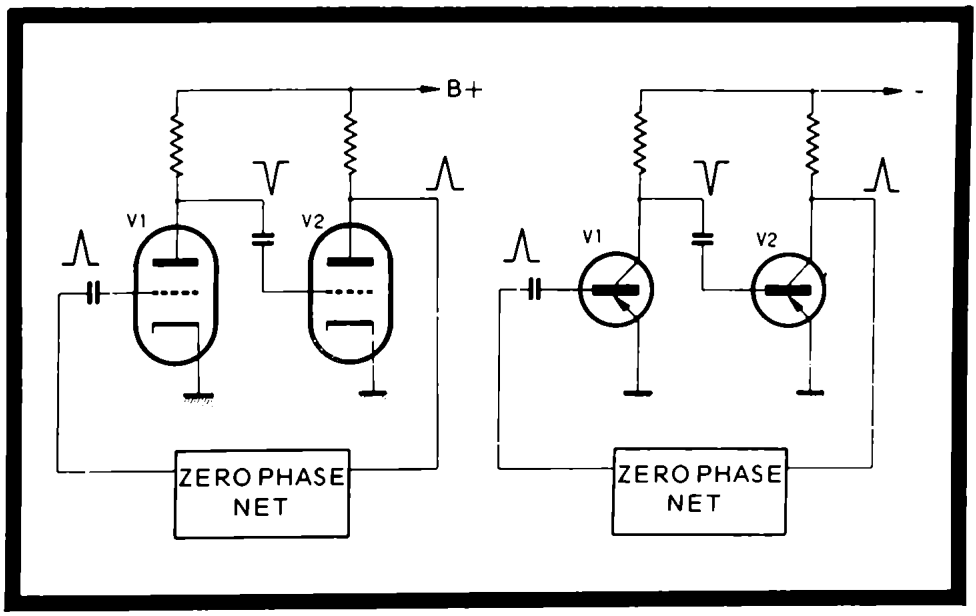


Fig. 3-45. A 2-stage amplifier needs no phase shift to oscillate.

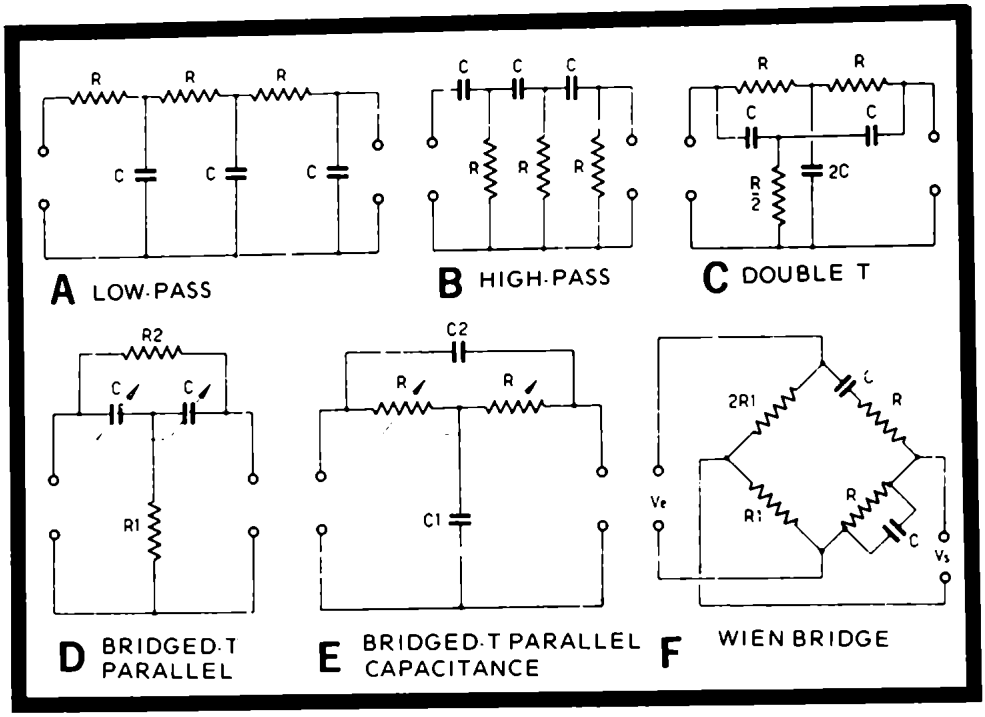


Fig. 3-46. Principal RC selective networks used in oscillator circuits.

low-pass (A) and high-pass (B) phase-shift networks, the double T (C), the parallel-resistance (D) and parallel-capacity (E), double T-derived bridged T networks, and the Wien bridge (f).

Introduced in the feedback loop of a back-coupled 2-stage amplifier, such as that in Fig. 3-45, the phase-shift networks produce oscillations at the frequency corresponding to an overall loop phase shift of 0 degrees, sufficient amplifier gain provided. This conditions is more significant here than formerly, for the 3-element RC networks in Fig. 3-46A and B introduce an attenuation of 29 which requires compensation. If amplifier gain is less than 29, sustained oscillations are no more possible, and an unstable selective amplifier results.

Since the RC networks are high-impedance devices, it is a good plan to have them feed an FET. This can be obtained with a common-source FET circuit (Fig. 3-47) featuring a gain well in excess of 18.4. (4-element RC networks, and as the one used, introduce an attenuation of 18.4 which must be made up.) Feedback potentiometer P2 adjusts the oscillation amplitude as well as the stability. This circuit was designed for a frequency as low as 10 Hz. The high-pass is good for low-frequency operation. For high frequencies, a low-pass is preferred. Bipolar transistor Q2 is used only as a current source to bias Q1. The FET is a P-channel type.

The double-T network (Fig. 3-46C) is used in precision oscillators yielding a high-purity sine wave, provided the three resistors and three capacitors have been adjusted to close tolerances. If the oscillator frequency is to be made variable, the double-T, which requires simultaneous variation of the three resistive or capacitive components, is sometimes replaced by the more convenient bridged-T networks (Figs. 3-46D and E) that need but two variable components, resistors or capacitors. At frequency

$$f = \frac{1}{2\pi RC}$$

with

$$C = \sqrt{C_1 C_2}$$

for Fig. 3-46E and

$$R = \sqrt{R_1 R_2}$$

for (Fig. 3-46D), the network impedance is practically infinite, and there is a zero input-output phase shift.

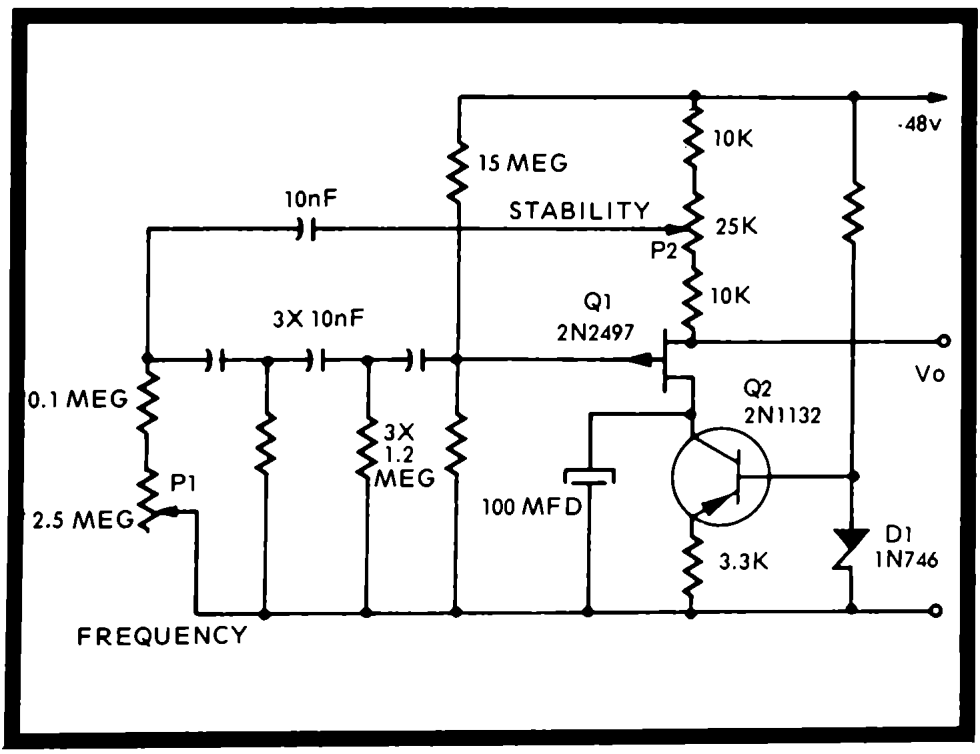


Fig. 3-47. RC phase-shift oscillator using an FET impedance converter.

The Wien bridge (Fig. 3-46F) is the selective network used in most variable-frequency generators of very low to high frequencies. Again, there is zero phase shift for:

$$f = \frac{1}{2\pi RC}$$

and a 2-stage amplifier is required to set up oscillations. But in contrast to the above selective networks, The Wien bridge has no terminal common to input and output and, therefore, has to be connected differently.

If a fixed frequency of high constancy is required, selective devices featuring very high Q-factors, such as mechanical resonators, are used. This may be a piezoelectric quartz crystal plate or a tuning fork, where oscillations are initiated and sustained by an electronic circuit. Frequency stabilities (df over f) of 10-5 to 10-6 are easily obtained using such devices, and 10-8 to 10-9 can be reached at least for short time intervals if supply voltages and operating temperatures are well stabilized (thermostats).

Crystal oscillators are serviceable between 50 kHz and 10 MHz and over. The tuning fork is conveniently used between 50 and 5,000 Hz.

Because of the inherently high Q-factor, oscillators using mechanical resonators and even LC circuits still yield an approximately sine-shaped wave when the oscillation amplitude is badly exaggerated. In contrast, RC oscillators produce a reasonably good sine-wave only as long as optimal operation conditions are closely approximated; as only slightly excessive amplitude will result in such a badly distorted waveform that the original sine wave can scarcely be guessed. Notice that the term "oscillator" generally implies "sine wave"; non-sine shapes are generated by wave (or waveform) generators such as the multivibrators.

MULTIVIBRATORS AND RELATED CIRCUITS

Among the non-sine-wave generators, **relaxation oscillators** produce pulses, square, triangular and sawtooth waves. Their repetition frequency depends on the RC time constant of a charging or discharging capacitor rather than on the frequency-selective properties of LC or RC circuits. Therefore, the frequency is much less stable and can be easily controlled over a large range by suitable **synchronizing** pulses.

Though invented back in 1919, when there was no electronics nor even radio yet, the **multivibrator** and its related

circuits are the foundation stones of modern electronics and computers, and it is important to understand them thoroughly and not to confuse them.

A multivibrator is a back-coupled 2-stage tube or transistor amplifier (Fig. 3-48) featuring a high loop gain. A small variation in any circuit voltage is immediately amplified and reinjected, so that rather violent oscillations build up, limited only by the cutoff and saturation conditions. This makes for a somewhat rectangular waveform. Actually, V1 and V2 are working alternately, the switch-over from one state to the other is transmitted by the RC coupling to the other stage which them immediately takes the complementary state. The repetition frequency of the multivibrator depends on time constants R1C1 and R2C2, the interstage couplings, and also on the electrode voltages or currents that determine the conducting and nonconducting states. By means of suitable sync pulses, the working cycle can be shortened to have the multivibrator operate at a higher repetition frequency.

If R1C1 is equal to R2C2, the multivibrator is symmetric, with a unity mark-space ratio. Different m-s ratios can be obtained with unequal time constants, though really narrow pulses cannot be generated by this circuit. To vary its repetition frequency while keeping the same m-s ratio, both resistances or capacitances must be varied simultaneously.

The circuit in Fig. 3-49 also exhibits one RC (or AC) and one DC coupling; however, **this is not a free-running multivibrator!** Look at the bias conditions: V2 is biased to conduct heavily, and V1 is practically nonconducting. This situation is further reinforced by the voltage drop across the common resistor due to the cumulative action described above. With no external signal applied, there is no reason for this state of things to change, and the circuit remains quiet.

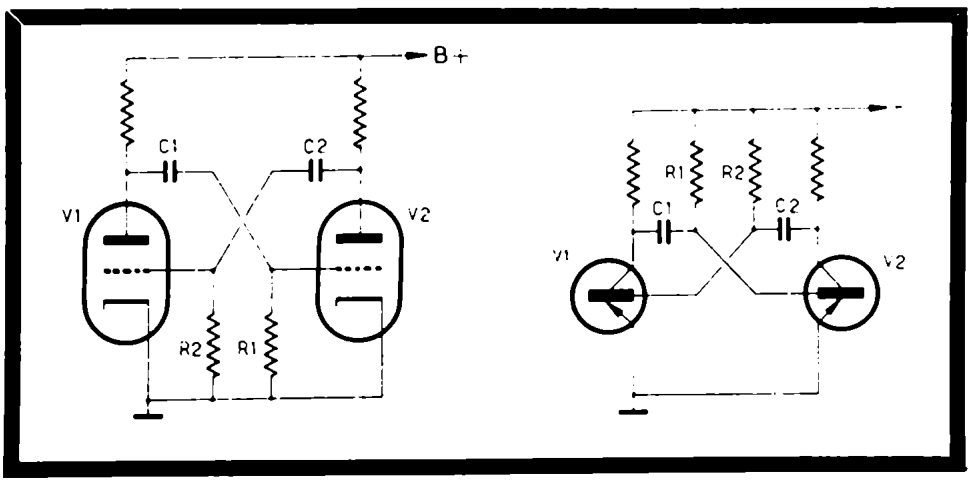


Fig. 3-48. Multivibrators use two capacitive couplings.

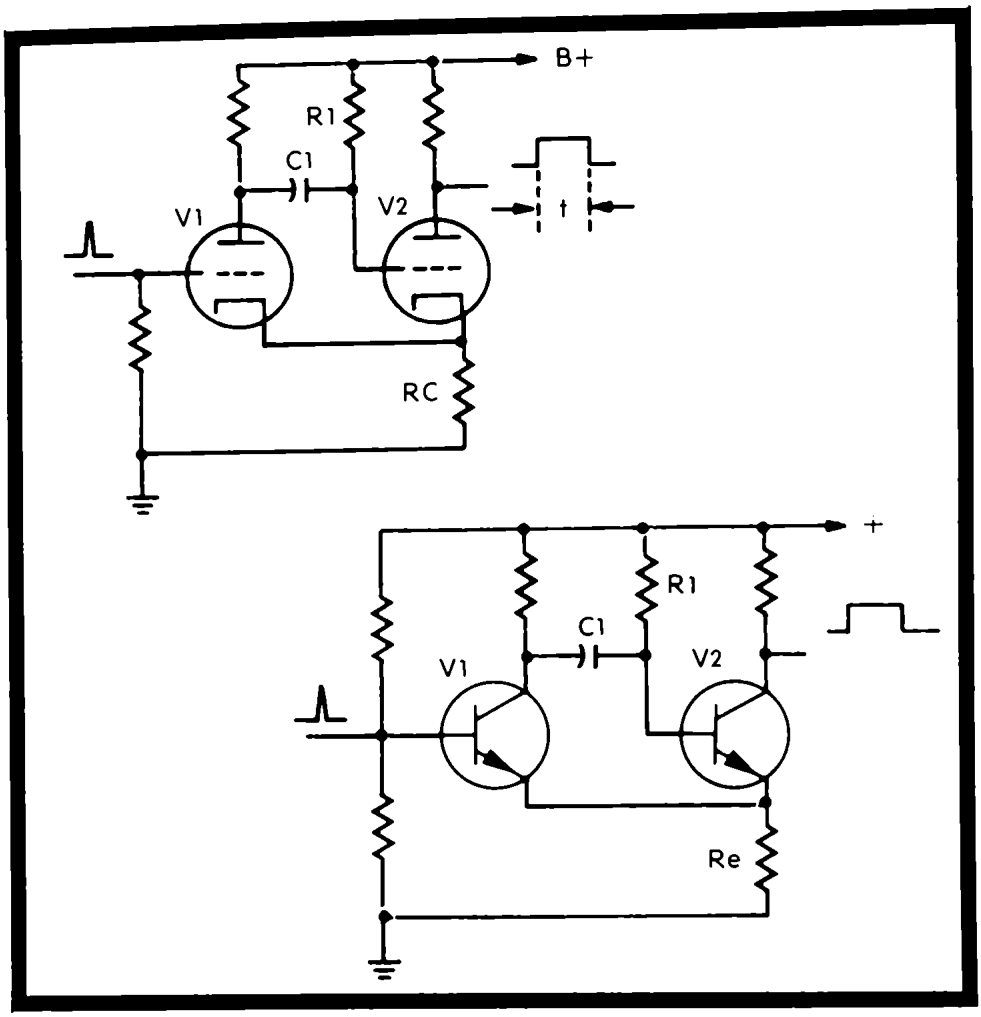


Fig. 3-49. Flip-flop circuits use one AC and one DC coupling.

But if we inject a positive-going pulse into the V1 input, this device will become temporarily conducting and transmit, via RC, a blocking pulse to V2. Thus, the operation of V2 will be toppled over, reversing the conduction situation. After a delay, depending on time constant RC, C will have discharged through R, then V2 recovers conduction and, by means of the voltage drop across the common resistor, throws V1 back to its initial nonconducting state. The circuit can be compared to a switch that first "flips over" and then "flops back," hence its name **flip-flop**. For the sake of clarity, we will stick to that term, though this circuit is also called single-shot multivibrator, or univibrator. An incoming pulse is transformed in a practically rectangular output pulse where the duration depends upon the RC time constant. The flip-flop, therefore, can be used to generate an adjustable time delay or to provide any incoming pulse with a well-defined duration.

The main characteristic of the flip-flop is its automatic return to a rest position, its only stable state, after a variable delay. This operating mode is realized when, in the quiescent state (no signal), one of the devices conducts while the other does not, a condition that can be created by different means.

Another minor modification leads to the **Schmitt trigger** (Fig. 3-50) comprising **two DC couplings**—by a common cathode or emitter resistor and by voltage divider R1R2 (capacitor C of small value is there only to speed flip-over; thus, it is not essential). In this circuit, V1 is conducting or not conducting according to the input electrode potentials, and V2 immediately takes the complementary state. Since the couplings are DC there is no reason for this situation to change with time—provided no signal reaches the V1 input.

The input electrode potential of V1 is set by potentiometer P. Rotating the control will throw V1 from conduction to nonconduction, whereupon V2 simultaneously takes the complementary (opposite) state. There is flip-over indeed, the feature common to all members of the multivibrator family, but it depends solely on the **level of the input voltage**. If a sine wave of sufficient amplitude is applied to the input terminals, the circuit will switch on and switch off when the sine-wave peaks two definite levels, V1 and V2. The Schmitt trigger is used to convert any waveform into a square wave and as a pulse shaper to give rounded pulses the neat square form needed by many pulse circuits.

Replacing capacitors C1 and C2 in the multivibrator in Fig. 3-48 with resistors R3 and R4 results in the **Eccles-Jordan** circuit in Fig. 3-51. This last member of the multivibrator

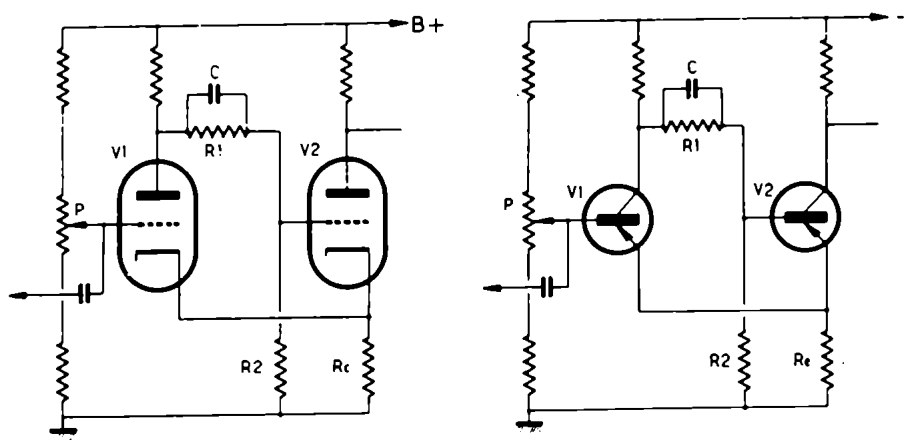


Fig. 3-50. A Schmitt trigger circuit uses two couplings.

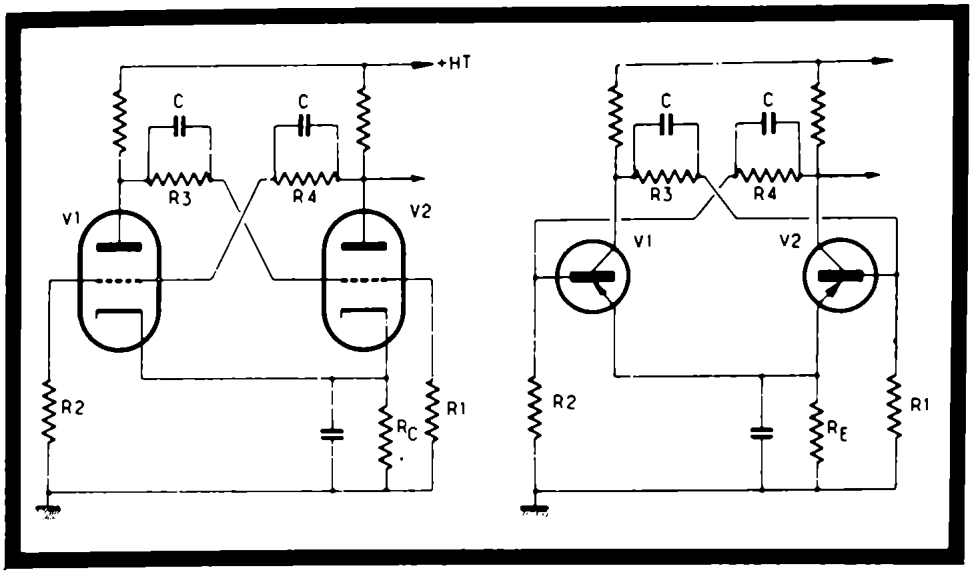


Fig. 3-51. An Eccles-Jordan or scale-of-two circuit also uses two DC couplings.

family also has **two DC couplings** and two stable states. One of the two devices is conducting at a time, the other is thrown off by the same cumulative effect as described above. Here again, no change occurs without a signal. Just as in the Schmitt trigger, the low-value capacitors (C) are used only to speed switching; although not fundamentally necessary, they may be required to allow reliable operation.

Suppose V1 is conducting to begin with; V2 then will be nonconducting. A suitable pulse applied to a suitable electrode (there are several possible injection points, due to signal polarity) will throw V1 out of conduction and make V2 conduct heavily. This is the second of the two states of the E-J circuit. Another pulse will throw it back to its former state, V1 conducting and V2 not. Thus, it takes two input pulses to obtain one output pulse (or rectangular half-wave) as shown by the oscillogram in Fig. 3-52 (the control pulses appear above the square-wave output). Hence, the name **scale-of-two** is given to the E-J circuit: it divides by two, and cascades of such circuits are used to scale-down generator frequencies.

The main application of E-J circuits is related to binary (digital) computing. The two device states, or the two possible output levels can be identified with the numbers 0 and 1, the basis of the **binary counting** system. The scale-of-two, therefore, is the basic computer circuit, and each arithmetic unit uses lots of them. Still another application made possible by the two states is the **memory** function, where an input is stored as a 0 or 1 output level. Table 3-2 sums up the main characteristics of various members of the multivibrator

family. (Caution: Some authors indifferently use the term "flip-flop" in reference not only to univibrators but also E-J circuits and sometimes even multivibrators, thereby leading to confusion of basically different functions. Thus, if you read somewhere of a flip-flop, first make sure what kind of circuit it really is.)

OTHER IMPORTANT CIRCUITS

The **blocking oscillator** (Fig. 3-53) is a tightly-coupled LC oscillator in which the tuning capacity is reduced to the stray capacitance of winding P (primary). This condition favors a

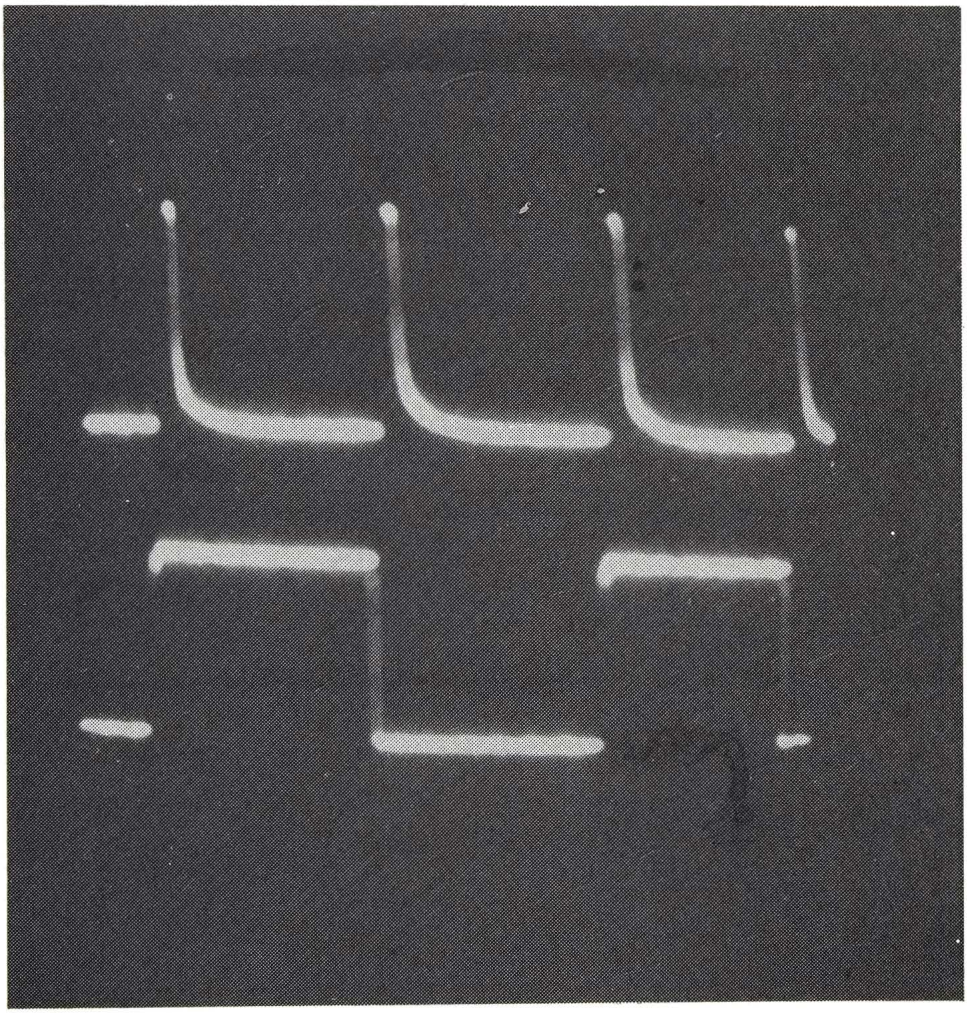


Fig. 3-52. Input (top) and output waveforms showing the performance of an Eccles-Jordan multivibrator.

Circuit type	Couplings	Properties
Multivibrator	2 AC	Unstable; free-running relaxation oscillator
Flip-flop or univibrator	1 AC, 1 DC	1 stable state; delay generator
Schmitt trigger	2 DC	2 stable states, switched by input voltage level; waveform shaping
Eccles-Jordan or scale-of-two circuit	2 DC	2 stable states, switched by pulses; counter and memory functions

Table 3-2. Multivibrator characteristics.

violent oscillation buildup so that after only one half-wave, capacitor C is sufficiently charged to block transistor operation. Instead of a sustained oscillation, the circuit generates only one half wave or pulse (diode D is often used to short-out the residual negative-going half wave, making for a "pure" positive pulse). After a time delay t, C discharges sufficiently through resistor R for the transistor to resume

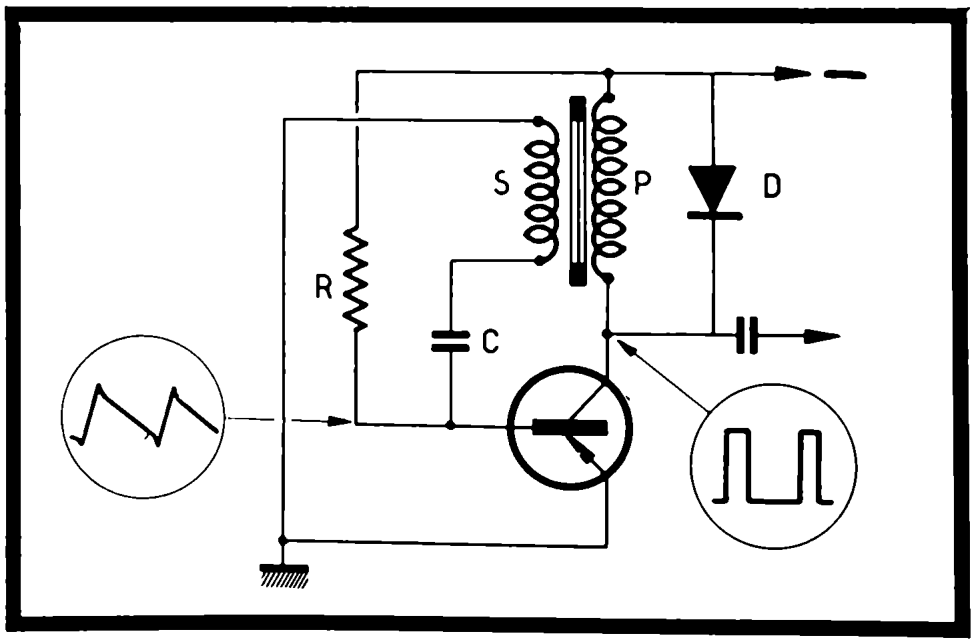


Fig. 3-53. A blocking oscillator generates sawtooth voltages and pulses.

conduction, setting up a new oscillation to be stopped after a sole half wave. Hence, the blocking oscillator is not an oscillator, but a large-amplitude **pulse generator.**

Since a blocking oscillator is a relaxation type somewhat like the multivibrator, its repetition frequency largely depends on time constant RC. Its inherent instability turns to an advantage, for it easily synchronizes on a wave or pulse train of a higher repetition frequency. This rapid charge and slow discharge of capacitor C results in a sawtooth wave on the base, usable for linear sweep applications. Returning R to a positive potential changes the circuit from the free-running to the triggered mode: the oscillator, normally quiescent, is triggered by a negative pulse applied to the base. After generating a unique pulse, the circuit comes to rest again, waiting for a new triggering pulse.

The sawtooth voltage (Vc) across C, however, is exponential rather than linear, as illustrated by the diagram in Fig. 3-54; the initial steep rise eventually flattens to reach charging voltage E very slowly (theoretically, this should take an infinite time). When a linear (i.e., constant-speed) sweep voltage is required, the "brute" sawtooth wave must be linearized by some constant-current charging or discharging circuit.

In electronic equipment, it is often necessary to switch on or to flash a pilot light by electronic control. While electroluminescent diodes are fully compatible with semiconductor circuitry, incandescent lamps generally consume too much power for integrated and possibly even transistor cir-

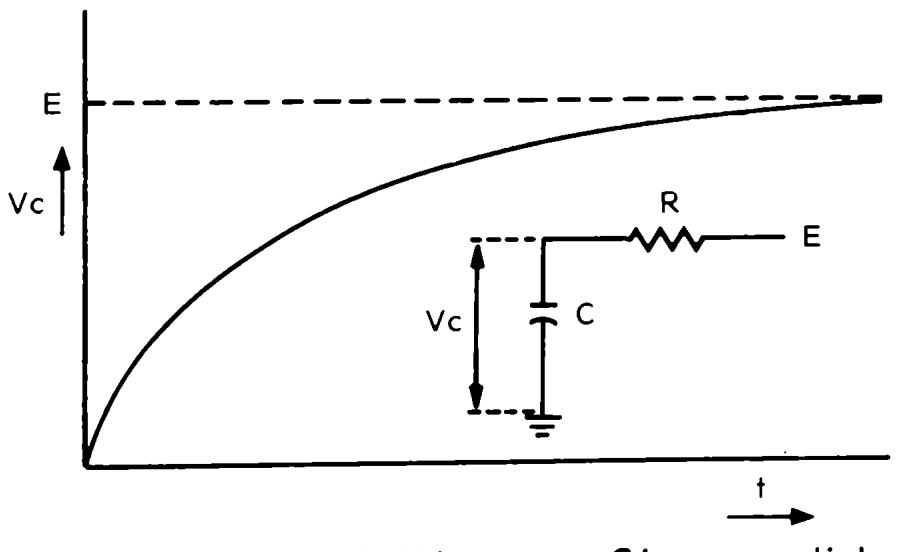


Fig. 3-54. The voltage buildup across C is exponential.

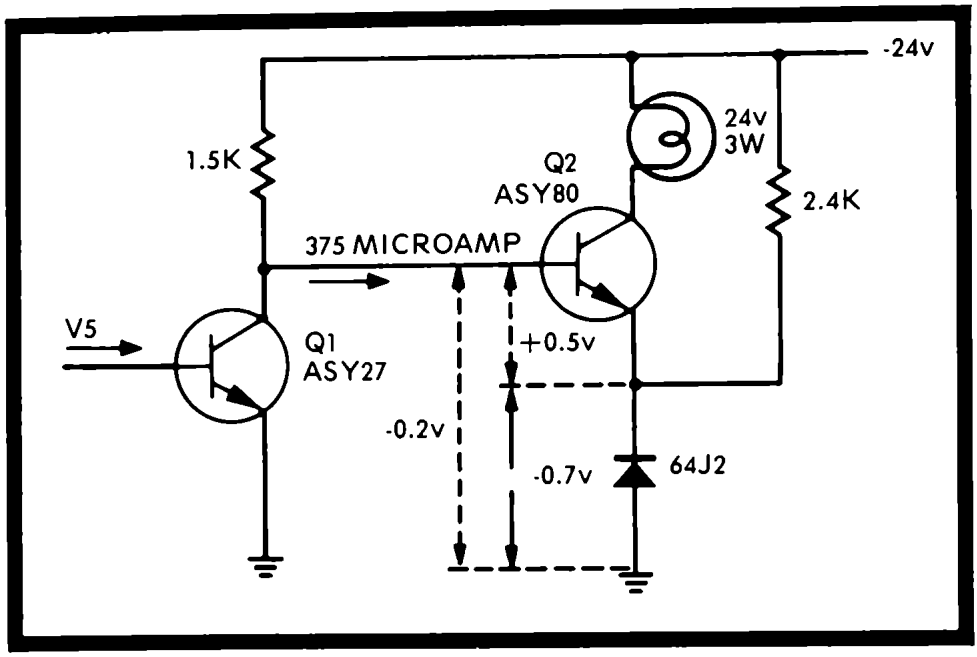


Fig. 3-55. Two-transistor lamp driver circuit for digital electronics.

cuitry. The 2-transistor **lamp driver** in Fig. 3-55 offers a solution to this problem. A 24v 3w lamp is used here; it has a steady-state current drain of 125 ma. When cold, the filament resistance is ten times lower; therefore, the transient lighting current reaches 1,250 ma. But since the thermal time constant of such a filament is rather small, the duration of this transient can be estimated at less than 50 millisec, allowing for the use of a transistor such as type ASY 27 with a specified 600 ma maximum peak collector current. This output transistor (Q2) is emitter-biased to -0.7v by a silicon diode; its base is directly tied to the collector of saturated driver transistor Q1 (Vce sat equals 0.2v). Thus, Q2 is normally nonconducting, and a current of 375 microampere applied to its base throws it into conduction and lights the lamp. As explained in the paragraph dealing with transistor switching, the transition from one (conduction) state to the other must be made instantly if Q2 is to survive; this calls for ample signal voltage Vs on the base of Q1.

PHASING CIRCUITS

It is often necessary to vary the phase of a sine-wave voltage in regard to a reference phase, especially for thyratron or thyristor control. Generally, passive RC circuits of the basic **differentiating** (D) and **integrating** (I) type are used (Fig. 3-56).

The operation of phase shifters is easily demonstrated by means of a double-trace oscilloscope. Fig. 3-57 shows the oscillogram of two sine waves of the same frequency in-phase, and approximate quadrature (90 degrees) is represented in Fig. 3-58 (the top of one wave corresponds to the zero-crossing of the other). The waves in Fig. 3-59 are out of phase, since their amplitudes are continually opposite.

The phase bridge in Fig. 3-60, using ganged potentiometers, yields two voltages, Y1 and Y2, phase-shifted in regard to X, but at a constant mutual out-of-phase condition. Such a device is useful for push-pull control of thyratrons and thyristors.

While all these phase shifters provide for phase variation, phase inverters form a special class of devices yielding rigid in-phase and out-of-phase signals. This can be done with a grounded center-tapped transformer fed input voltage Vi and yielding out-of-phase output voltages Vo1 and Vo2 (Fig. 3-61A). But such a transformer is too bulky for ICs, so the apparently more complicated solution in Fig. 3-61B, using a transistor phase inverter, may be preferred. The transistor operates linearily for low distortion. Collector voltage V01 is out-of-phase in respect to emitter voltage V02 and input voltage Vi. These outputs can be used to drive a push-pull stage.

TRANSFORMATION OF WAVEFORMS

A sine wave injected into the differentiating or integrating circuits (D and I) in Fig. 3-56 will be attenuated and leave the

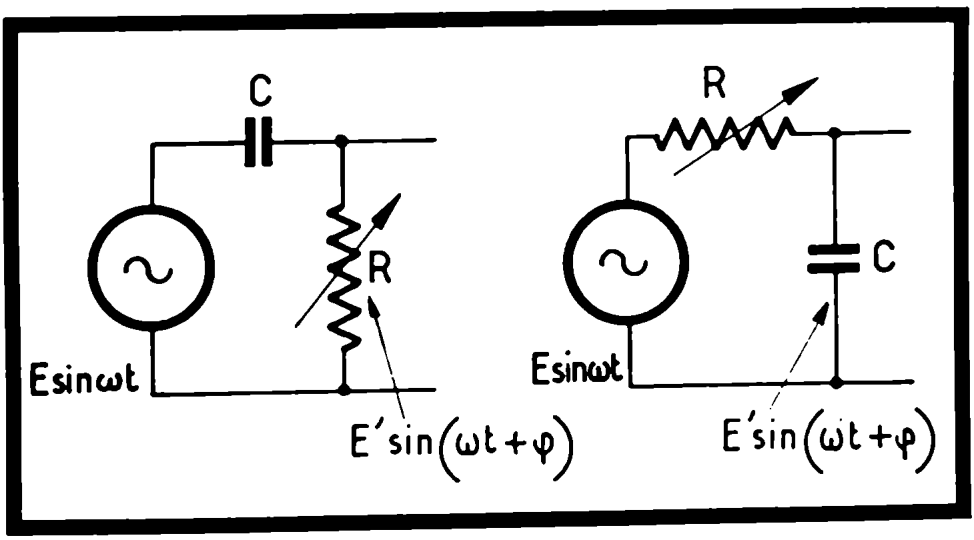


Fig. 3-56. Differentiating (D) and (I) integrating circuits (right).

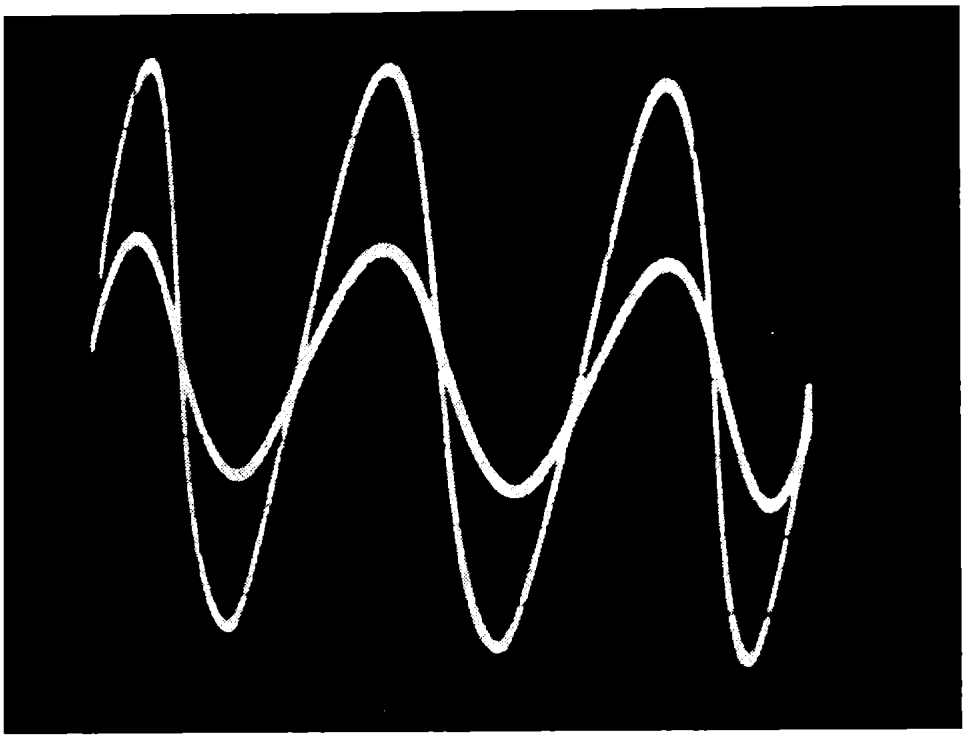


Fig. 3-57. Two sine waves of the same frequency and phase.

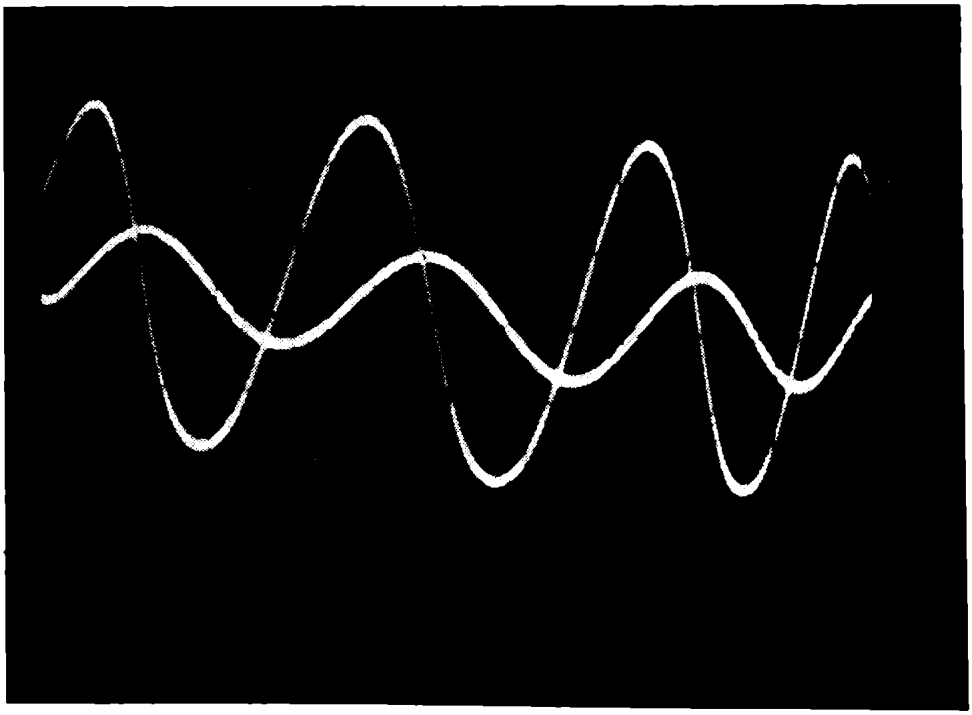


Fig. 3-58. These sine waves are approximately 90 degrees apart.

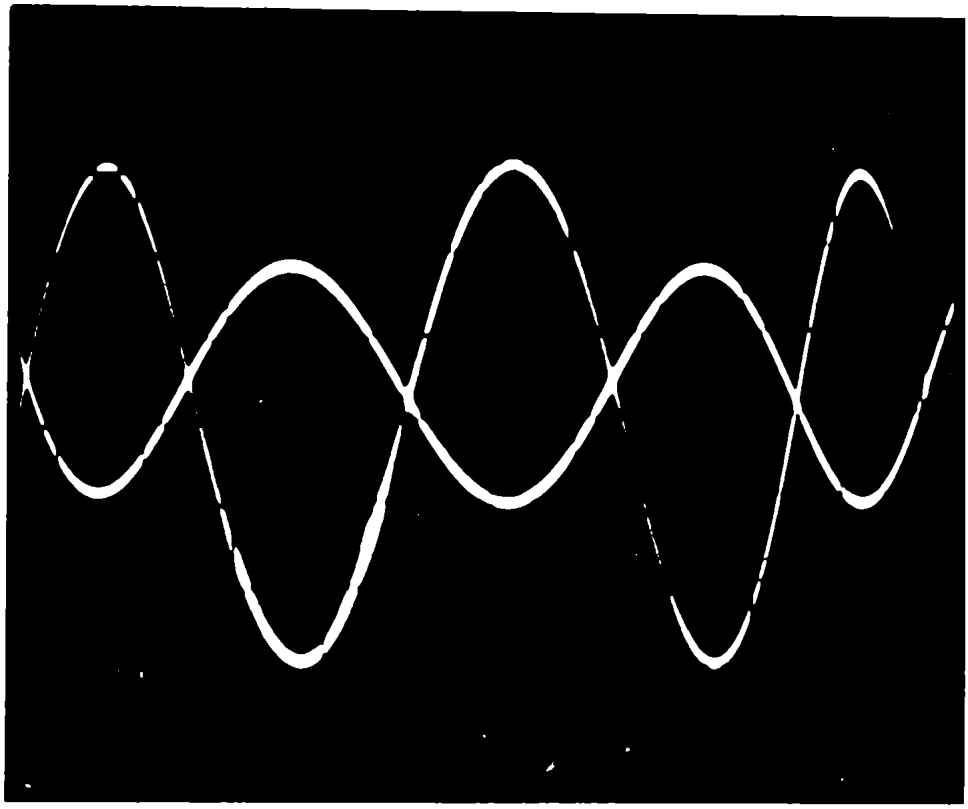


Fig. 3-59. Sine waves completely out of phase (180 degrees).

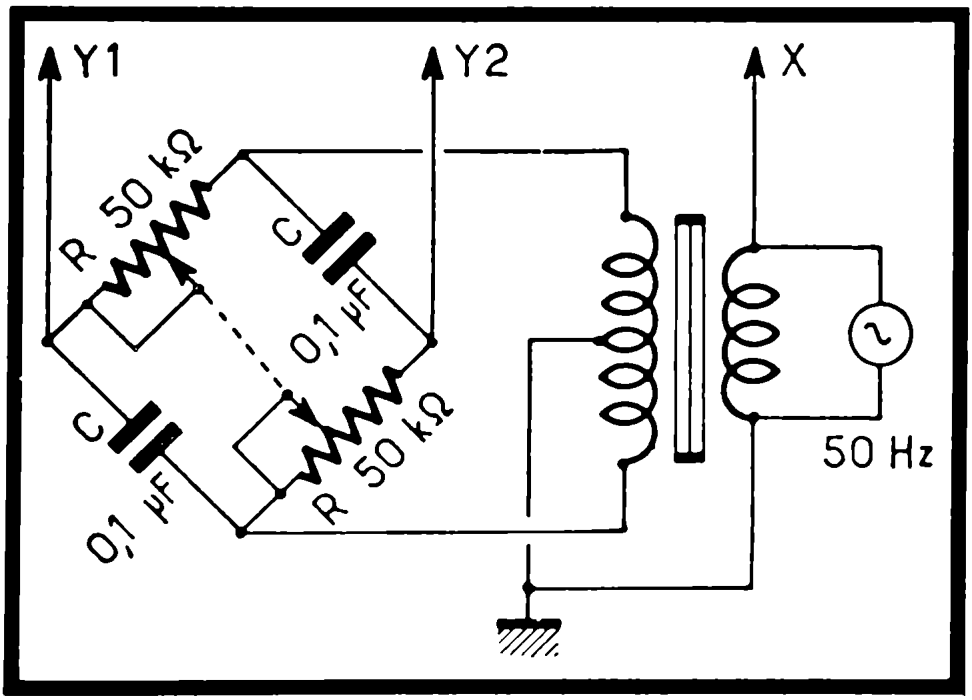


Fig. 3-60. Phase bridge providing two voltages continually out of phase.

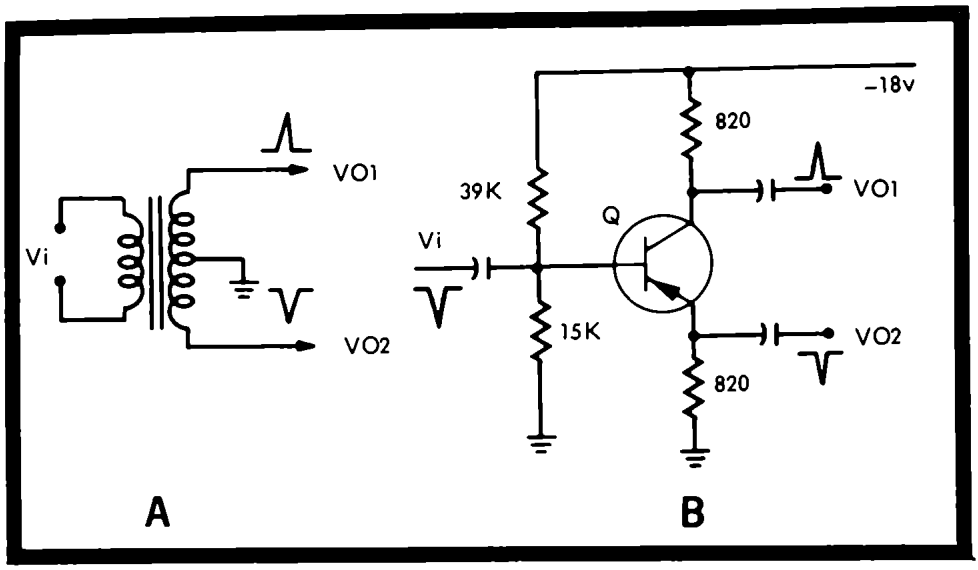


Fig. 3-61. Transistor and transformer phase inverter circuits.

circuit at a different phase; however, its shape, will remain unchanged. The action of these circuits actually is frequency-dependent, so that waveforms containing more than one frequency component will be distorted by unequal attenuation of these components. Sine waves comprising but one component are not altered. For the same reason, pulses cannot be phased directly without badly distorting them. So a good plan may be to start with an easily produced and phased sine wave used as a "raw material" for pulse generation.

Differentiating and integrating circuits owe their names to the mathematical operations of differentiation and integration they carry out on the transmitted waveforms. **Differentiation** results in an exaggeration of the rapidly changing waveform components without disturbing the slowly varying parts. Of an applied square wave, the D-circuit retains the steep rising and falling slopes, but eliminates the flat top and bottom elements, so that a series of alternate positive and negative pulses results (Fig. 3-62). The unequally-spaced pulses were obtained using a square wave with a non-unity mark-space ratio.

Integration, the opposite operation, results in ironing-out rapid waveform changes while integrating the general tendency of the variation. Thus, integrating a square wave yields a triangular wave.

The differentiating and integrating efficiency of D and I depends on the relation between time constant RC and duration T of a period of the processed wave. It can be shown (mathematically and oscilloscopically) that good differentiation requires RC to be equal to or less than 0.01T, while

good integration is realized with RC to be equal to or greater than T.

A more or less sine wave can be squared by clipping its positive and negative summits, for instance, by means of two back-to-back connected zeners. But you have to clip a major part of the wave to obtain an approximately square-looking truncated wave. The oscillogram in Fig. 3-63 shows the trapezoid wave obtained with insufficient input voltage (top), and a better-looking approximately square waveform obtained with five times the former voltage (bottom). But even this last wave has rounded tops and corners. A well-squared wave can be generated only by the Schmitt trigger in Fig. 3-50. Complex waves such as television signals are composed of pulses and rectangular waves generated individually and then added in suitable mixers. To keep the different wave components phase-rigidity, all generators used must be mutually synchronized.

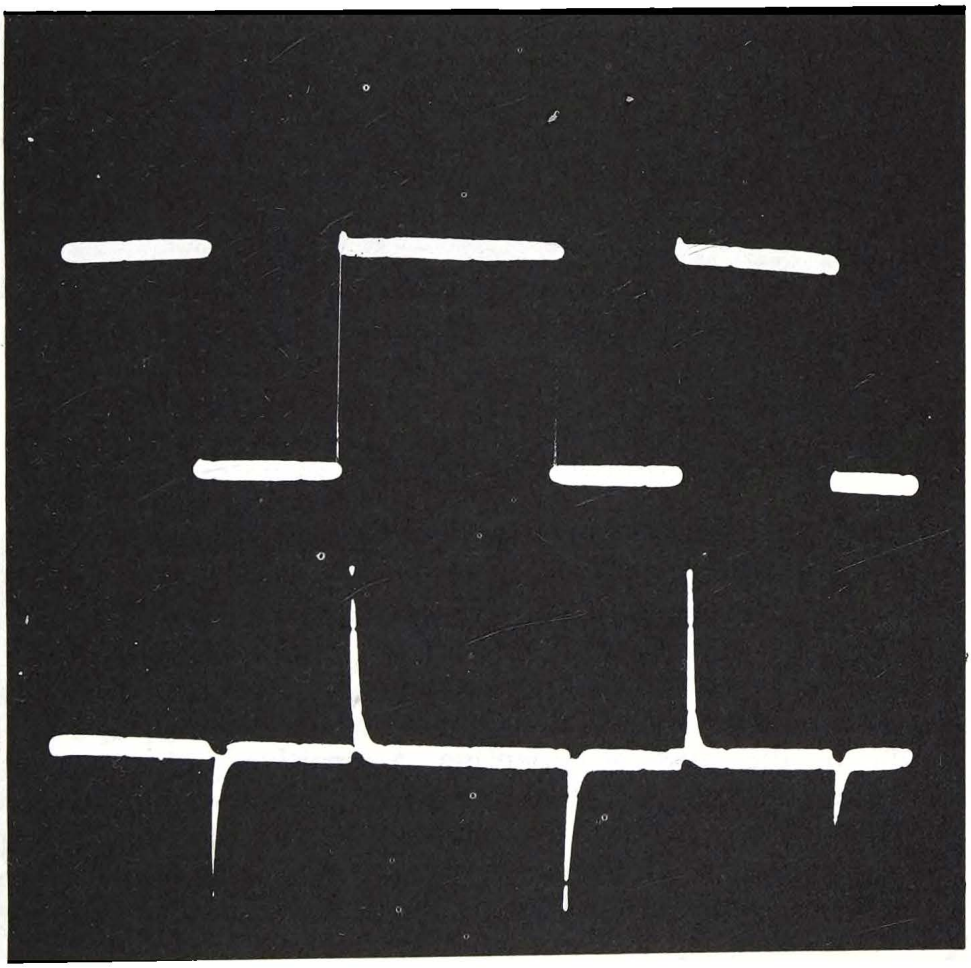


Fig. 3-62. A differentiated non-unity mark-space ratio square wave results in unequally spaced pulses.

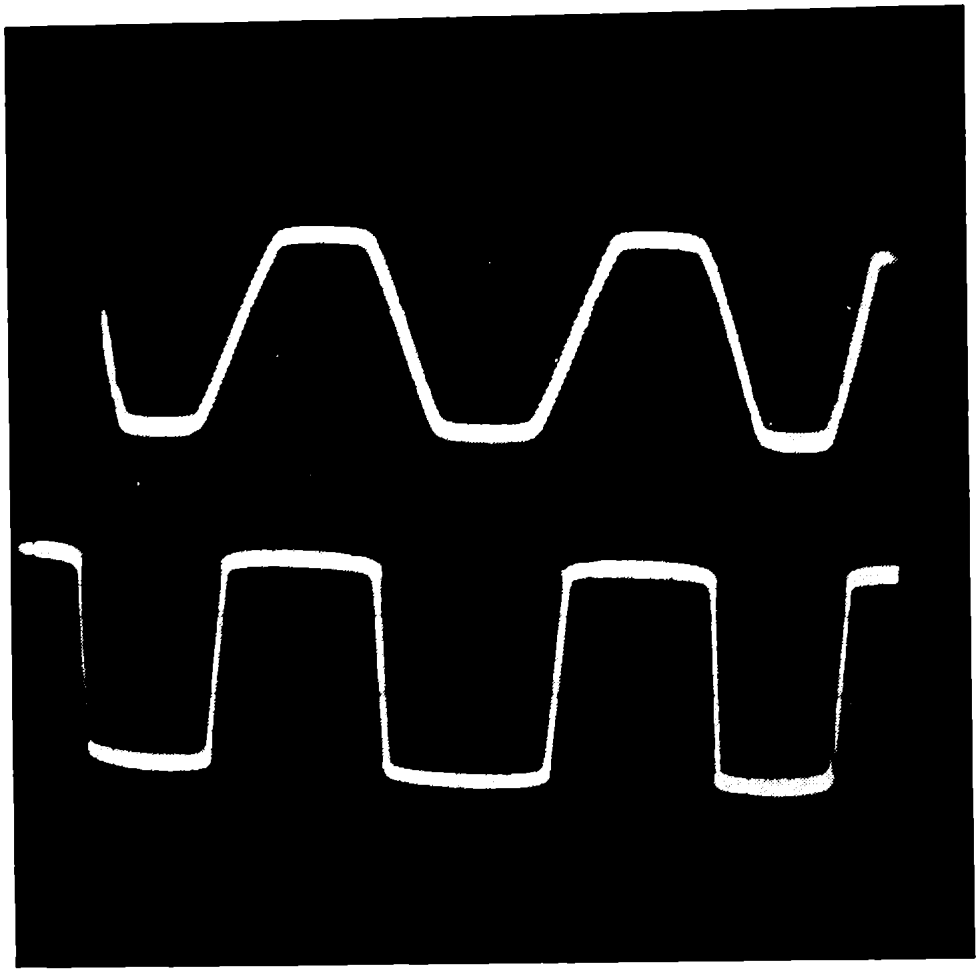


Fig. 3-63. Clipped sine waves indicate insufficient (top) and adequate driving voltage.

HOT-CATHODE THYRATRONS

A thyratron is a triode tube containing some rare gas which radically modifies its behavior. The thyratron, indeed, is an on-off device having but two states—conducting or not. Normally, its grid is negatively biased, and if we apply a B+ voltage to its anode, nothing happens. The thyratron stands the voltage applied and there is no anode current in load RL. Reducing the bias or injecting a positive pulse causes ionization ("fires") of the thyratron gas (a faint glow is generally visible through the glass bulb), and tube resistance falls to quite a low value, so low that an excessive current can build up and destroy the tube if an adequate current-limiting resistance is not providied.

Thus the grid initiates the tube firing, but once the tube is fired, the grid looses control and conduction will carry-on indefinitely until the anode supply is interrupted. With suf-

ficient bias provided, the device will remain in the non-conducting state until a new pulse or bias reduction strikes it again. Thus, to de-ionize the gas tube, the anode voltage is lowered (at least momentarily) to only several volts, a level called the extinction voltage. There evidently is a problem in DC-fed circuits, but it can be overcome by shortcuts such as the RC-network in Fig. 3-64. This is not a filter, but a short-time power supply. Capacitor C charges through high-value resistor R, and the charged capacitor supplies power to the thyratron when it strikes. However, this will discharge C more rapidly than it can recharge. So the voltage across C will fall and eventually become sufficiently low to de-ionize the tube. With AC-fed thyratrons, there is no problem, for the tube ceases conduction at the end of every positive half cycle. If the firing signal is still present, the tube will strike again at the following positive half cycle.

For small power controls (relays, fractional HP motors), tetrode thyratrons such as types 2050 or 2D21-PL21 are preferred. The second grid (unlike that in vacuum tetrodes) is generally tied to the cathode like a suppressor grid. A tetrode has a much lower grid current than a triode thyratron, and grid leaks up to 10 meg are possible, making for a welcome high-impedance input. The cathode current is higher, too. But even more important is the much higher **control ratio**, or ratio of anode voltage to triggering bias. Depending on the working conditions, this control ratio ranges between 100 and 500 for the 2D21, resulting in a trigger sensitivity much higher than that of a triode thyratron.

The **electronic timer** in Fig. 3-65 is an application of a 2D21. It generates accurate time delays between half a second and several minutes. The shortcoming of such timers, based

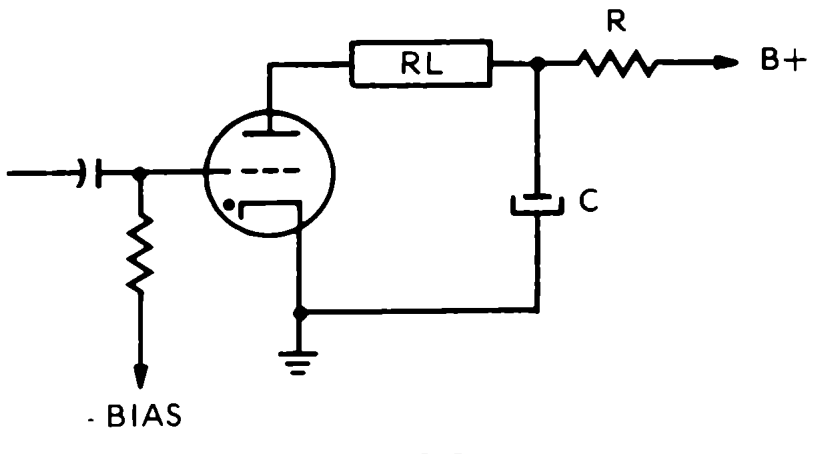


Fig. 3-64. Self-quenching DC-fed thyratron circuit.

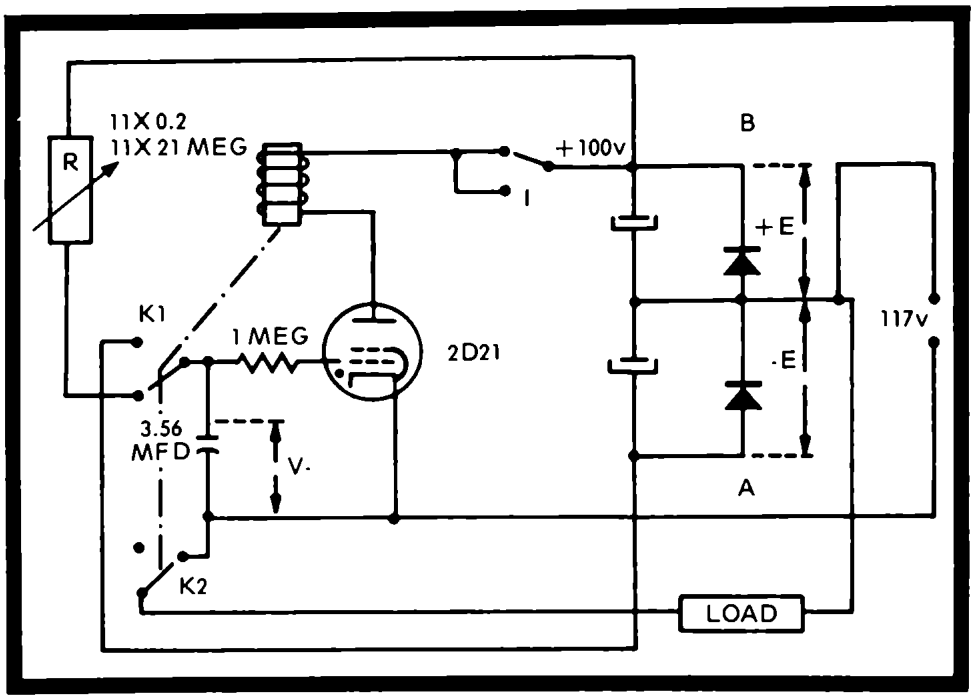


Fig. 3-65. Precision timer circuit using a small tetrode thyratron.

on the charge or discharge of a capacitor, generally is the dependence of the time delay on the level of the charging voltage, so that stabilization is mandatory. In the proposed timer, capacitor C is successively charged by two opposing sources of the same voltage, +E and -E. By thyratron-controlled relay contact K1, C is charged first to -100v (Point A) and then discharged through calibrated variable resistor R, connected to +100v (Point B). It can be shown that time t necessary for voltage v across C to pass from -100 to 0v is given by t equals 0.69 RC (with t in seconds, R in megohms and C in microfarads), independent of the actual level of E.

Some instants after a cold start, with C tied to +100v by K1, the thyratron fires, energizes the relay and connects C to +100v via R. With the thyratron ionized, the grid, now negative, has lost control, and it is necessary to momentarily interrupt the anode current acting upon double-throw toggle switch I. This de-energizes the thyratron, the relay armature falls back, and C charges through R, which is connected to +100v. After a time delay t, v equals 0, the thyratron fires again and ends the delay. The load is switched by contact K2.

COLD-CATHODE THYRATRONS

Among the small-power devices, hot-cathode thyratrons have gradually been superseded by cold-cathode types, for a

heaterless tube obviously has higher reliability and service life expectancy; it dissipates no power when idling and has no start-up delay. On the other hand, its maximum anode voltage Vmax is limited to about 330v, its ionization voltage, Vi, is rather high (about 100v, against about 10v for the hot-cathode variety), making for a considerable voltage drop, so it requires a somewhat high control voltage (about 50v), a problem if the control signal is extremely small.

This device, indeed, is a glow tube with an additional flat doughnut-shape control electrode (G) inserted between a straight-wire anode and a large-surface cathode (Fig. 3-66). The working anode voltage has to be lower than Vmax if the tube is not to fire without any gate control, but in excess of Vi for ionization to be possible. With the tube extinguished, a positive pulse applied to gate G ionizes the gate-cathode space, thereby triggering the main cathode-anode discharge. From then on, the gate looses control. The ignited tube offers a low resistance, and if sufficient series resistance is not provided for current-drain limiting, the current density in some cathode spots may rise enough to initiate thermal emission. The glow discharge then becomes a destructive flashover.

The triggering characteristic (Fig. 3-67) describes the operation of the cold-cathode thyratron. Although the closed-

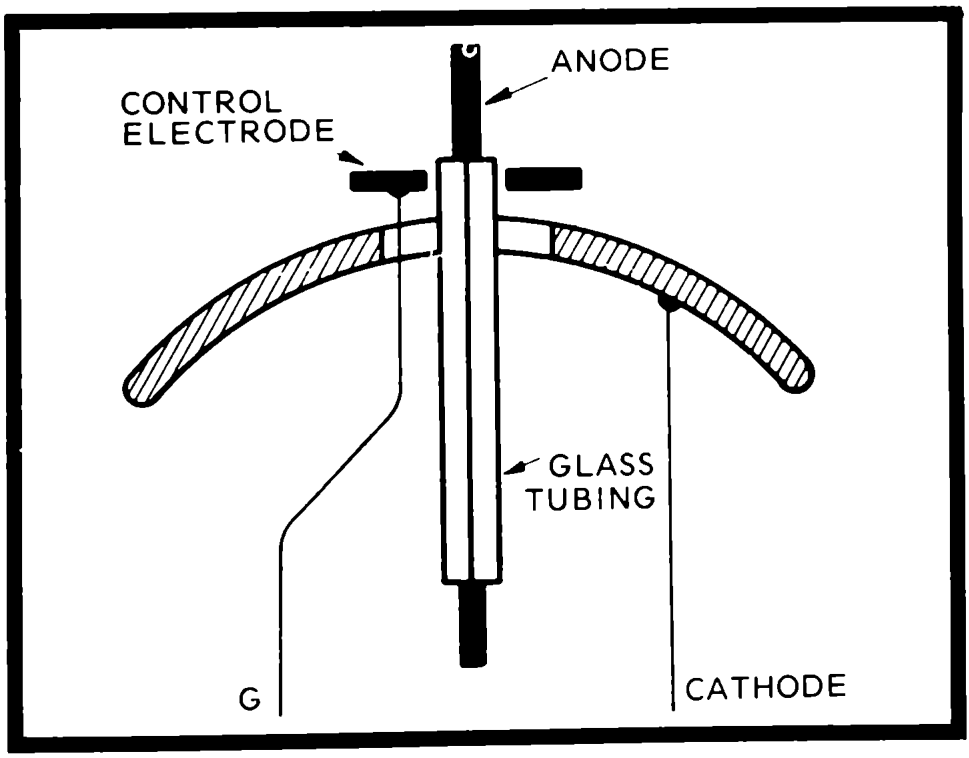


Fig. 3-66. Drawing of the electrode arrangement in a cold-cathode thyratron.

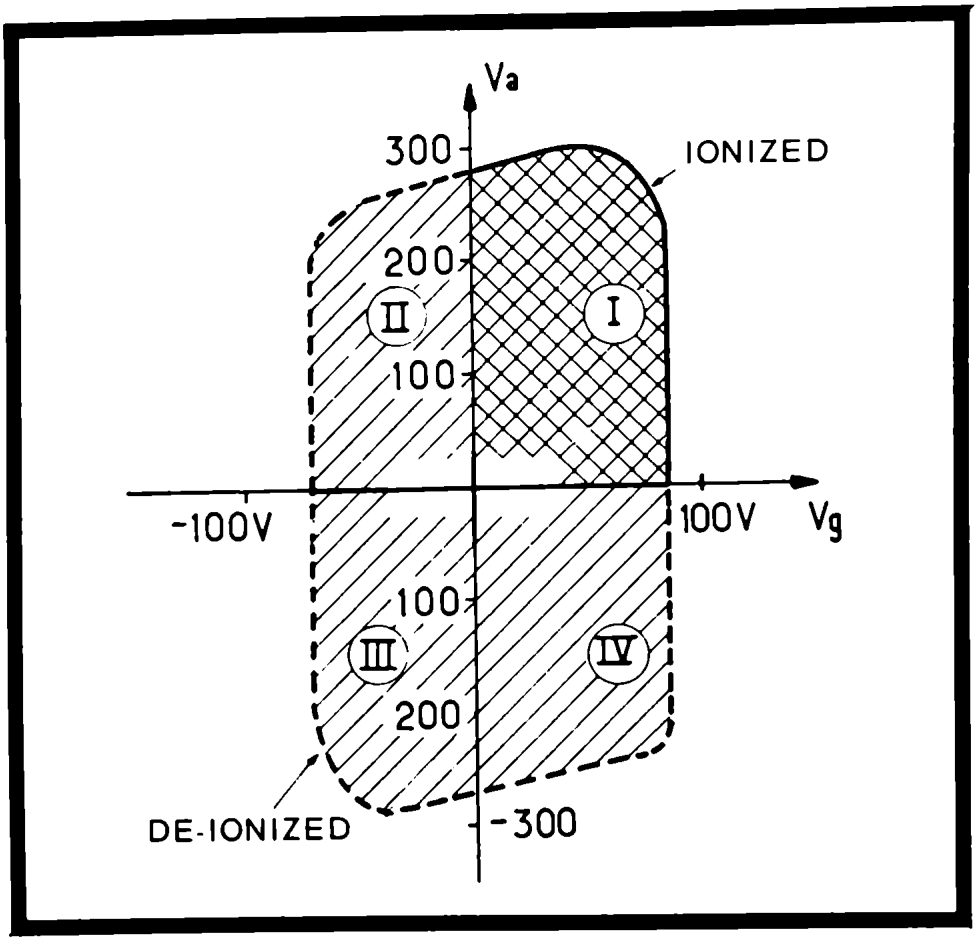


Fig. 3-67. Triggering characteristics of a cold-cathode thyratron.

loop working diagram covers the four quadrants I to IV, only I is recommended for normal operation with positive anode (Va) and gate (Vg) voltages. The tube is extinguished if its working point is situated in the region enclosed by the diagram, and ignited outside of this region. On the diagram shown, if Va is less than 300v, an input voltage Vg greater than 80v is required to strike the tube, and this value is independent of Va. To reduce this fairly high control voltage requirement and increase the device sensitivity, the gate can be biased to, say, 70v, so that a pulse of about 10v is sufficient for firing.

Cold-cathode thyratrons can be DC- as well as AC-fed. In the photoelectric control in Fig. 3-68, the gate is biased by a voltage divider R1R2. If R1 is made a photoconductor, its resistance is reduced by illumination, making the gate more positive and firing the tube, thus energizing the relay. (If the relay is to operate when illumination ceases, R1 and R2 are permuted.) Resistor R charges small capacitor C (not more

than about 5 nf), so that upon firing, C discharges into the gate and thus prevents gate current Ig from loading the voltage divider.

If this circuit is AC-fed, ionization ceases at the end of every positive half wave, and re-ignition occurs only if the gate still is sufficiently positive. If DC-fed, the thyratron can be de-energized only by momentarily reducing its anode voltage below Vi by interrupting the B+ supply or by applying a large negative voltage pulse on the anode.

THYRISTORS AND OTHER SOLID-STATE SWITCHING DEVICES

The 4-layer diode is a 2-terminal semiconductor device comprised of three junctions between regions P1N1, N1P2 and P2N2. If a voltage of the polarity shown in Fig. 3-69 is applied to its terminals, junctions P1N1 and P2N2 are forward-connected to conduct. The central junction, N1P2, however, is reverse-connected and it blocks current. Hence, the device is nonconducting.

If now we gradually increase the applied voltage (VB) by means of potentiometer P, for a given voltage (VBO of, say, 80v, avalanche multiplication takes place, and the leakage current increases rapidly, putting the device in the high conduction region where it will stay as long as anode current I

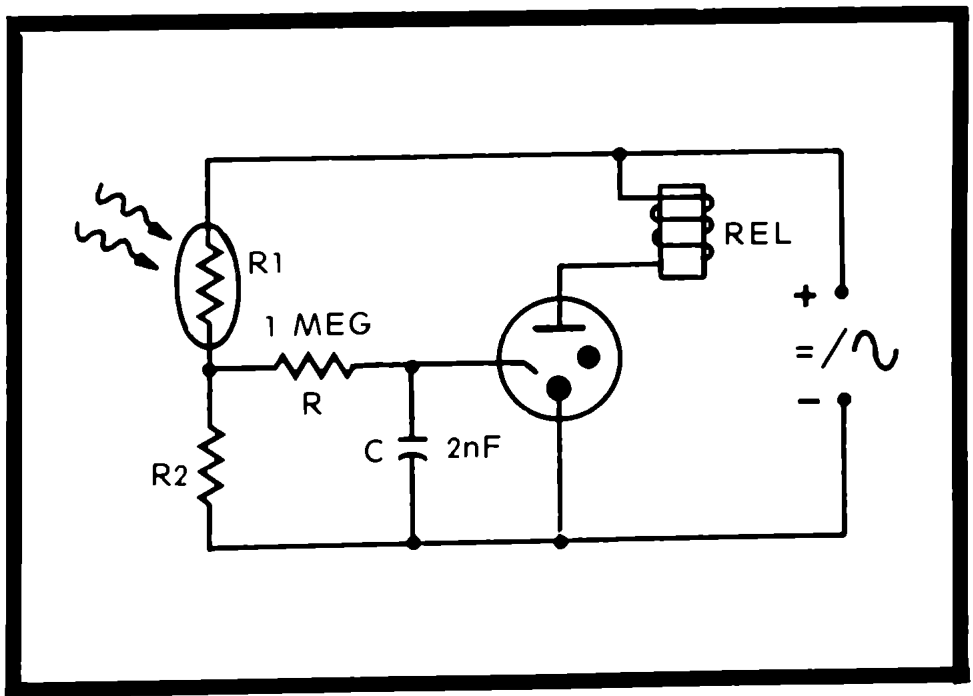
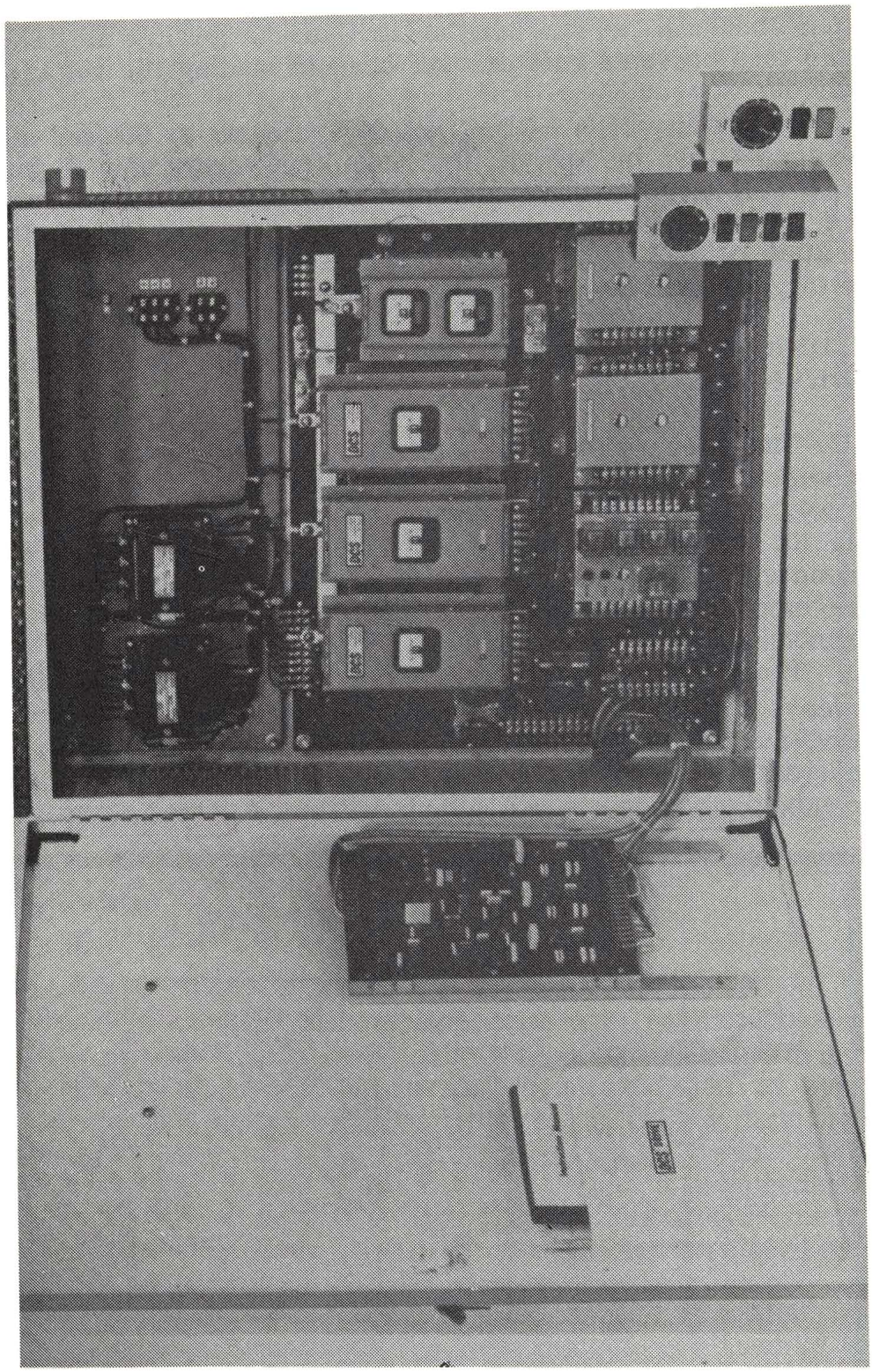


Fig. 3-68. Photoelectric relay control circuit using a cold cathode thyratron.



Thyristors are used in this motor controller which is designed to continuously vary the speed of DC motors rated from 7½ to 200 HP. Manufactured by Robicon Corp.

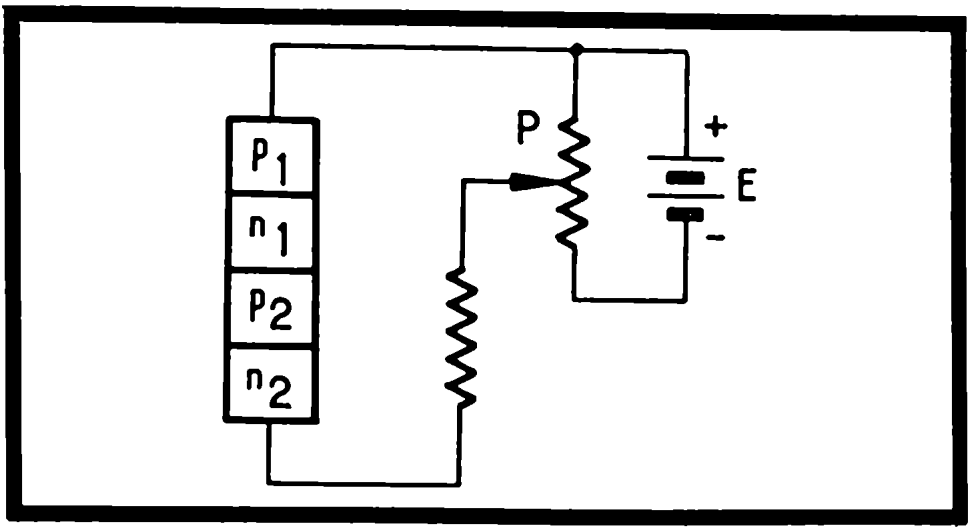


Fig. 3-69. Sketch showing the arrangement of the N and P regions in a 4-layer diode.

at least equals holding current Io (Fig. 3-70). The device immediately will revert to its forward blocking state if I drops below Io. VBO is called the **forward breakover voltage**. The left-hand reverse characteristic is similar to that of ordinary silicon rectifiers; a reverse blocking region ends with the

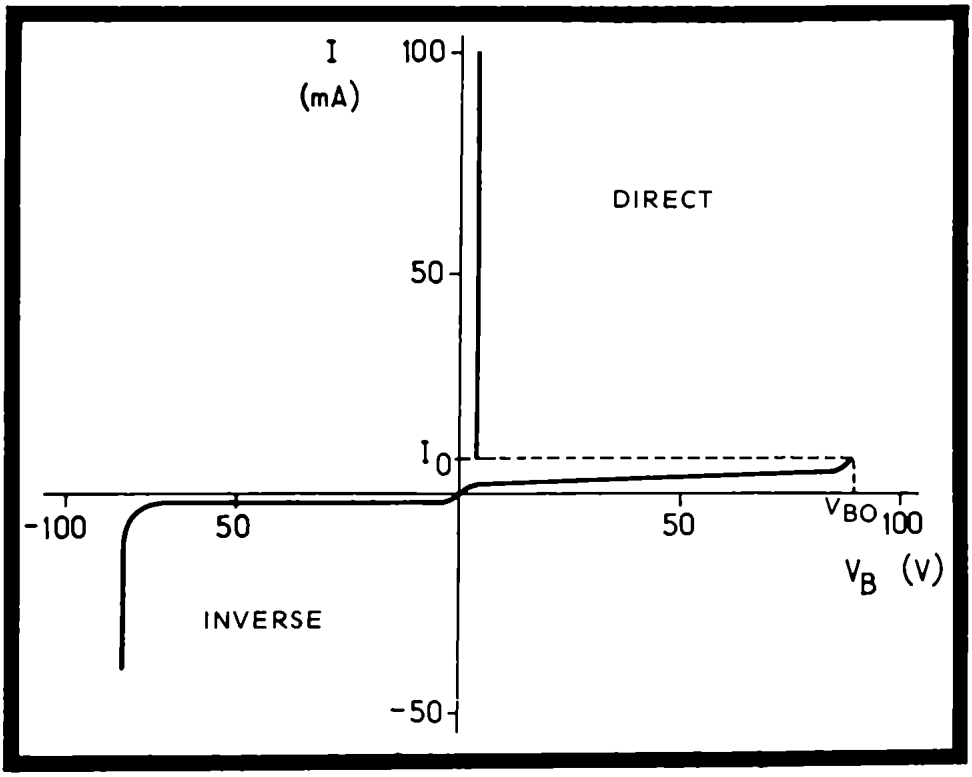


Fig. 3-70. Current (I) voltage (VB) characteristic of a 4-layer diode.

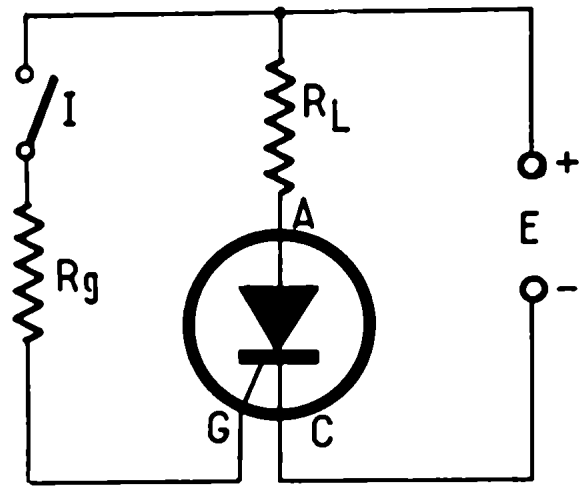


Fig. 3-71. Basic DC-fed thyristor circuit.

maximum reverse voltage point where the reverse avalanche conduction region starts. Most P-N-P-N devices are designed with a peak reverse voltage capability at least equal in magnitude to the forward breakover voltage.

The flexibility of this device is greatly increased by connecting one or more of the intermediate layers N1 and-or P2 to a terminal to be used as a control electrode or gate G. The 3-electrode device obtained is called a **4-layer switch, thyristor or silicon controlled rectifier** (SCR)

Now consider the fundamental circuit in Fig. 3-71. If switch I is open, there is no gate current and the device behaves like a 4-layer diode. To put it into conduction, a voltage at least equal to VBO is applied, and the characteristic BAOCDEF in Fig. 3-72, corresponding to the unique diagram in Fig. 3-70, is obtained. If now we close I to have G pass current Ig, the breakover occurs at a voltage where V1 is less than VBO, and the forward characteristic changes to OC'-D'E'F (the reverse characteristic, of course, remains unchanged). Hence, increasing Ig reduces the forward breakover voltage and shifts Point D to the left. For a given value of Ig, the forward blocking region disappears altogether, and a normal rectifier characteristic, such as indicated by the dotted line, results.

This peculiar characteristic explains the use of the thyristor as a "solid-state thyratron." Consider a supply voltage VO between VBO and V1. With Ig equal to 0 (switch I open), the working point is H; the thyristor operates in its forward blocking region and passes but a small leakage

current. If I is closed, a gate current develops, and with Ig sufficiently positive, the working point changes from H to D' and, this being a breakover point, from D' to the branch E'F in the high-conduction region where it settles at a point corresponding to the load current drain. If this current exceeds the rating, the junctions overheat and the thyristor is destroyed. As heating takes time, a very short strong current transient may be better supported than a much lower continuous DC current.

Just as in a thyratron, the gate looses control once the thyristor has been fired. To return it to the blocking state, the anode current must be reduced momentarily below holding current Io (corresponding to the ionization voltage Vi of the thyratron). In AC-fed circuits, this automatically occurs at the end of every positive half-wave, and conduction resumes only at the next positive half wave if there is sufficient gate current (corresponding to the thyratron grid voltage). In DC-fed circuits, there are several ways to stop conduction: the anode current can be interrupted or the device can be shorted momentarily, or a suitably injected pulse can reduce the anode current below Io.

Thyristors are serviceable in applications covering an extremely wide variety, their massive use in electronic controls for industry (and in appliances, too) is just starting

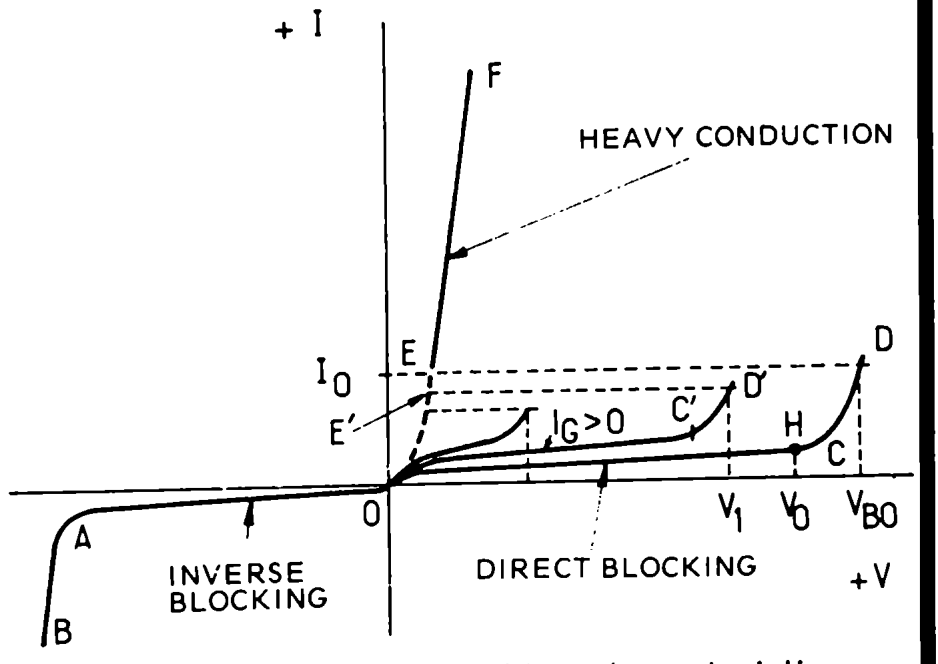


Fig. 3-72. Transistor switching characteristics.

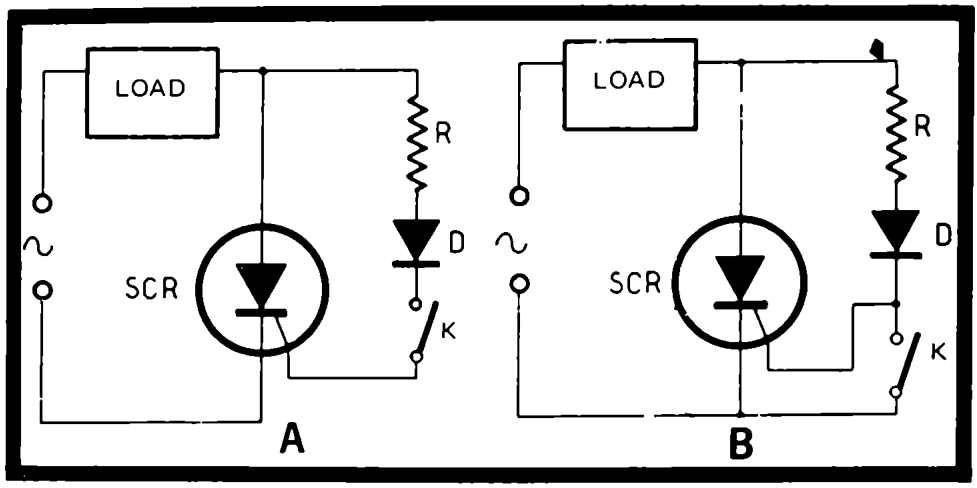


Fig. 3-73. Two methods of controlling an AC-fed thyristor (SCR).

now. Fundamental circuits for AC supply are indicated in Fig. 3-73. (There are commercial thyristors rated for a direct 117 or 220v mains connection, so that a step-down transformer can be saved.) The thyristor is used as a static power relay to switch a series-connected load, RL. Devices designed to pass hundreds of amps while standing more than one thousand volts are marketed. Two such devices directly fed by a 480-volt (RMS) line are able to control a 300-kilowatt load!

In Fig. 3-73A, the thyristor is fired when switch K is closed. Gate current Ig is limited by series resistor R. The device passes current only during the positive half waves, and diode D will not inject a reverse voltage into G during the half waves of nonconduction. (This diode also can be connected between gate and cathode.) When K is opened, the device, already de-energized during the negative half cycles, will not strike again at the next positive half cycle. In Fig. 3-73B, switch K' (normally closed) ties the gate to the cathode, hence Ig equals 0, just as in Fig. 3-73A, with switch K open. Now K' must be open to fire the thyristor.

Because of the very low control current (Ig), a thyristor can be controlled by a low-power switch such as a sensitive relay or a contact-closing thermometer, or a static control resistive transducer (photoconductor, thermistor) can be used as well.

The circuits described are on-off controls. In AC-fed thyristor circuits, a stepless variation of controlled power, however, is easily obtained by displacing the thyristor firing point along the time axis during the positive half cycle (de-energizing takes place anyway toward the end of that half cycle). By delaying the firing point, a smaller portion of the half wave is used and less power is delivered to the load. This

allows for controlling the speed of an electric motor or dimming a light. While these applications are discussed in Chapter 8, the simple stepless half-wave power control in Fig. 3-74 is typical for such applications. At the beginning of the positive half-wave, capacitor C charges through variable resistor R, and the thyristor anode voltage rises simultaneously. After a time delay, mainly depending on time constant RC, a current flows through diode D1 into the gate and fires the device. The thyristor remains conducting up to about the end of the half cycle. The 50K variable resistor allows for a phase shift of nearly 180 degrees, making for a gradual control of rectified power from about zero to full power. During the negative half cycle, C charges in the opposite direction and blocks the thyristor at the beginning of the next positive half cycle.

The universal motor-speed drive in Fig. 3-75 is just another application of this principle. Universal or AC-DC motors are used in applications requiring high torque during starting and at low speeds (the armature and field are series-connected). In the present circuit, thyristor SCR is inserted between them, allowing half-wave operation. The degree of power rectified and the motor speed depend on the SCR firing angle during the positive half cycle as formerly, and is a function of time constant RC. Variable resistor R allows for speed adjustment. Voltage across RC is stabilized by zener diode DZ during positive half-cycles (diode D1) to avoid the effect of mains fluctuations on the speed setting.

Stabilization of the preset speed is obtained by applying the counter-electromotive force developed across the rotating

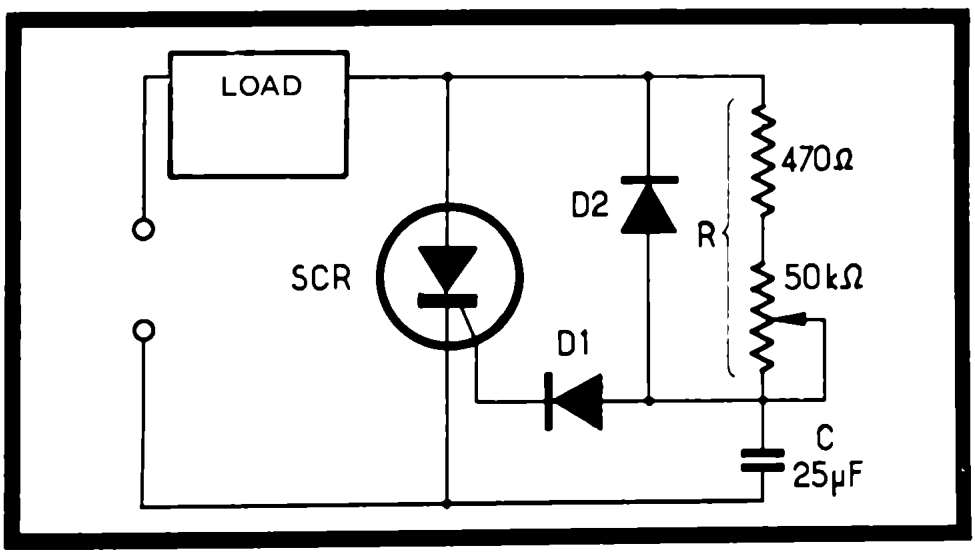


Fig. 3-74. Variable half-wave load current control circuit, useful as a light dimmer or motor control.

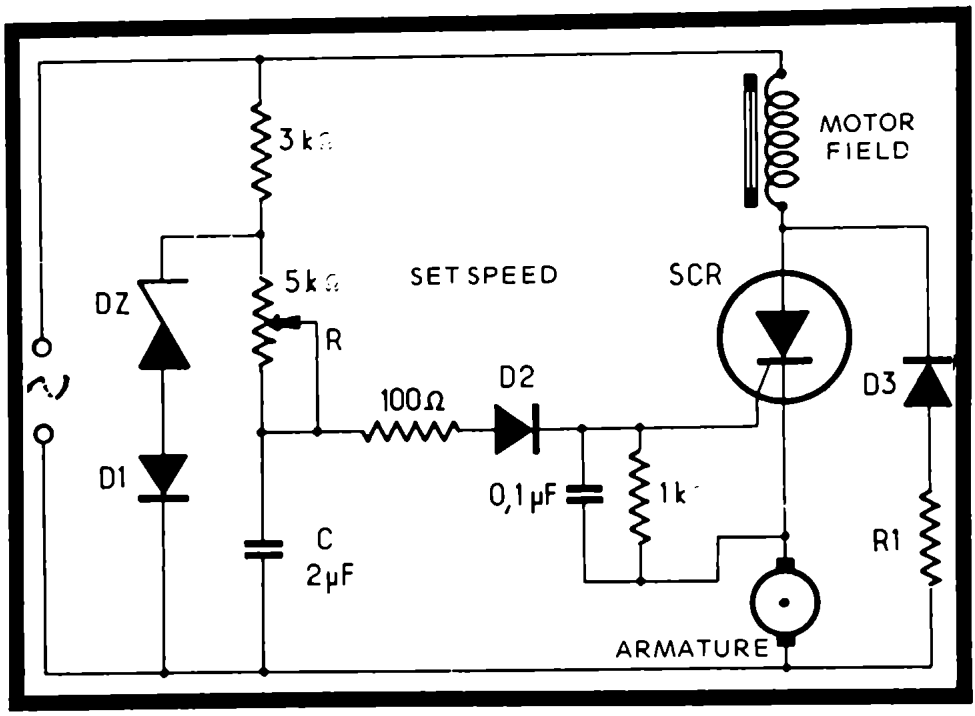


Fig. 3-75. Half-wave series motor speed control circuit.

armature to capacitor C with a polarity opposing that of the charging current. Thus, motor acceleration has the same effect as a reduction in charging current and results in a slightly smaller amount of power delivery by the SCR, which slows down the motor. During the blocking state of the SCR, the field coil is fed through D3 and R1. (Though the universal motor is reversible by crossing armature or field connections, this stabilization principle works only for one direction of rotation.)

If a small glass window is fitted in the opaque SCR casing to let an external light impinge upon the light-sensitive junction, a **photothyristor** or light-activated silicon controlled rectifier (LASCR) results. Such a device can be triggered by a light ray as well as by an electric signal. An LASCR can replace a photoelectric transducer followed by a thyratron or a thyristor. Photothyristor applications are illustrated in Fig. 2-54 and 2-55, Chapter 2. While LASCRs are sensitive to visible light, their peak sensitivity is well within the infrared region of the radiation spectrum. If its rating is inadequate for energizing a power device, a LASCR can be used to drive a higher-power SCR.

The thyristor is a half-wave device; for full-wave control, two such devices are connected back-to-back. The Triac (General Electric), a full-wave switch with but one control gate, behaves as two back-to-back connected thyristors. This

3-terminal device with electrodes T1, T2 and G is triggered by a positive or negative control current flowing between G and T1. Its characteristic, shown in Fig. 3-76 is the now familiar thyristor characteristic symmetrically displayed in the first and third quadrants. This bidirectionality allows for considerable simplification of AC control circuits.

Another new component, the Diac (General Electric) or bidirectional diode, has been developed as a full-wave control device. Its characteristic (Fig. 3-77), displaying the same symmetry as that of the Triac, shows that up to a threshold voltage of VBO of some 35v positive as well as negative, the device is blocked and passes only a negligible residual current. Then, the breakover takes place at both polarities, and the current greatly increases while the voltage drop actually decreases, indicating a negative-resistance region (A and B). A variable full-wave load control, using a Triac triggered by a Diac, is shown in Fig. 3-78. Just as in the half-wave counterpart (Fig. 3-74) of this circuit, the firing angle is varied by a gate voltage phase-shift determined by RC. As soon as the rising slope of the positive as well as the negative half-cycle reaches breakover voltage threshold VBO, the Diac

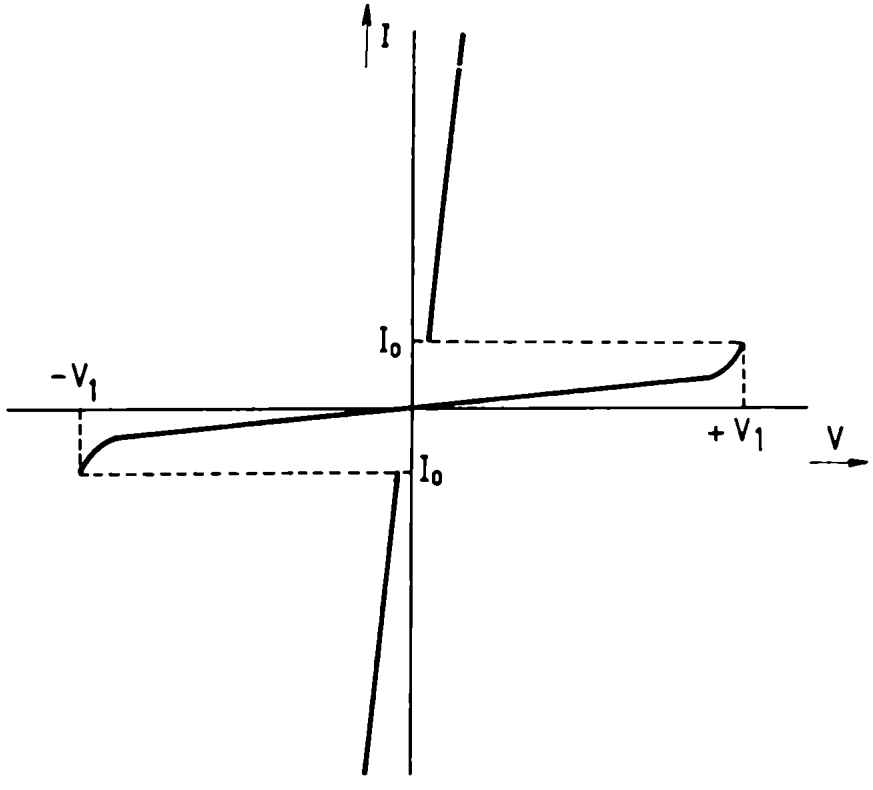


Fig. 3-76. Triac characteristic curve.

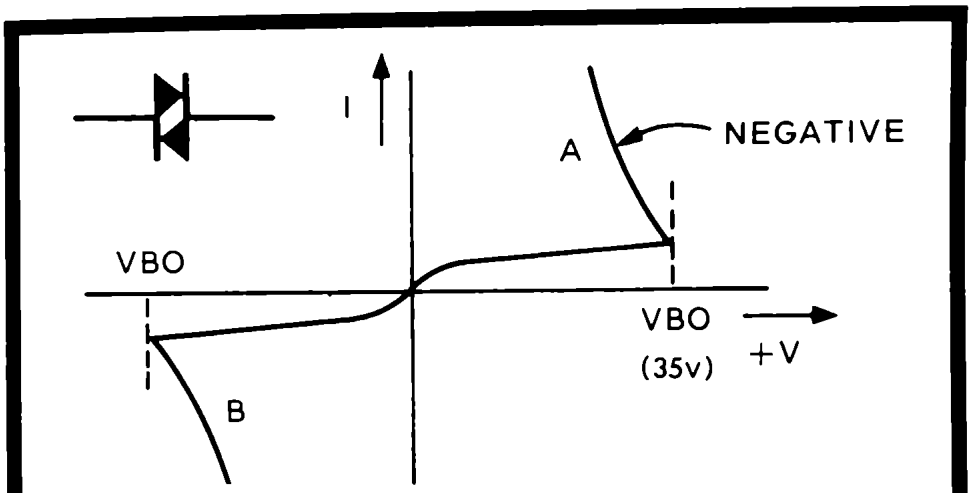


Fig. 3-77. Diac characteristic curve.

sharply begins conducting and transmits to the gate a steep voltage step that immediately fires the Triac.

THE UNIJUNCTION TRANSISTOR (UJT)

The little world of semiconductors definitely is rather complicated, for here comes the **unijunction transistor** (UJT), still another 3-terminal semiconductor device exhibiting very different characteristics. As shown in Fig. 3-79A, an UJT is made up of a tiny N-type silicon bar with electrodes called base-one (B1) and base-two (B2). Also, there is a small P-type region called the emitter (E), arranged in the center of the bar. The UJT symbol is shown in Fig. 3-79C and Fig. 3-79B is

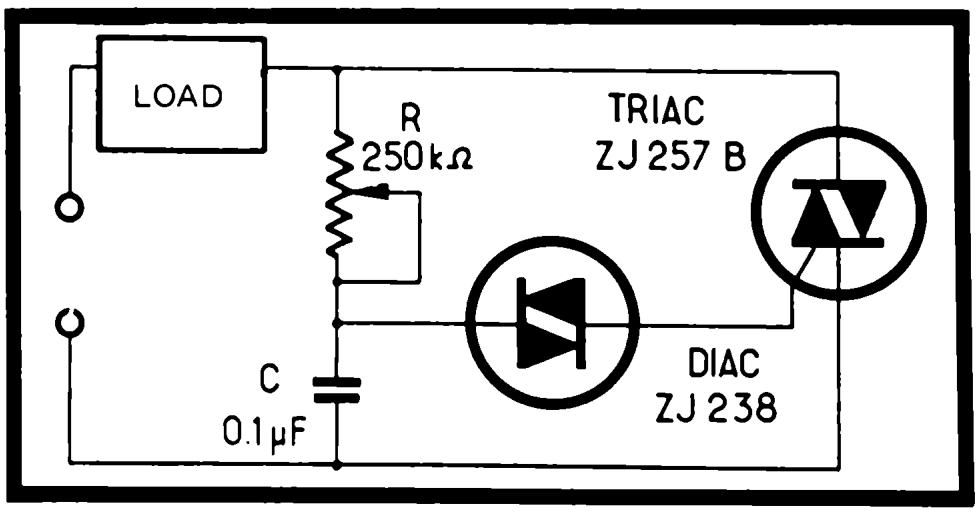


Fig. 3-78. Full-wave load control with phase adjustment.

its equivalent circuit, composed of voltage divider RB1 and RB2. Midpoint A is connected to a diode D.

B1 is common to input and output circuits. Positive voltages are applied to E and B2. As long as VA, the potential at Point A, is lower than emitter voltage VE, the diode is blocked, and the emitter current IE is 0. This condition makes for a comparatively high value of RB1 of, say, 4,600 ohms. If VA is greater than VE, the diode begins conducting, a current passes from E to B1, and RB1 decreases. If IE is 50 ma, RB1 should be 40 ohms. The main UJT applications is as a relaxation oscillator, especially for thyristor triggering.

The fundamental UJT relaxation oscillator circuit diagram appears in Fig. 3-80. Capacitor C charges through resistor R, and VE rises gradually. When VE equals VA, the junction begins conducting, the resistance of RB2 decreases, and C rapidly discharges through base-one resistor B1. At the end of the discharge, VE increases again, and a new relaxation cycle sets in. The waveforms on the electrodes are indicated on the diagram—an approximate sawtooth wave on E and pulses of opposite polarity on B1 and B2. The repetition frequency depends on time constant RC. This circuit is used as a trigger pulse generator, as a timer and as a time base (sawtooth linearization provided). It can be synchronized by injecting negative pulses into E.

A simple AC-fed thyristor control triggered by a UJT is shown in Fig. 3-81. Capacitor C charges through variable resistor R and fires the UJT which delivers a triggering pulse to the SCR gate. The pulse strikes the SCR, switching it to

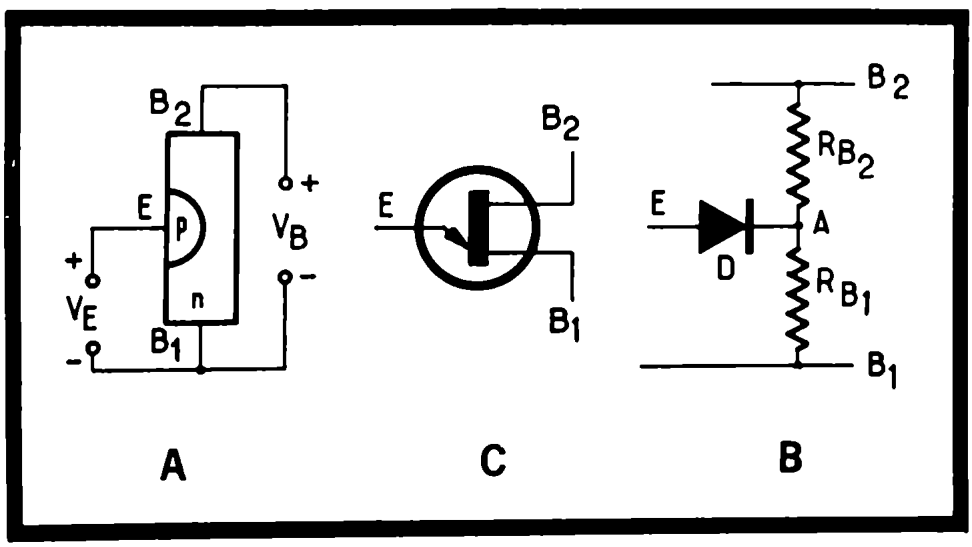


Fig. 3-79. Structural diagram, equivalent circuit and symbol of a unijunction transistor (UJT).

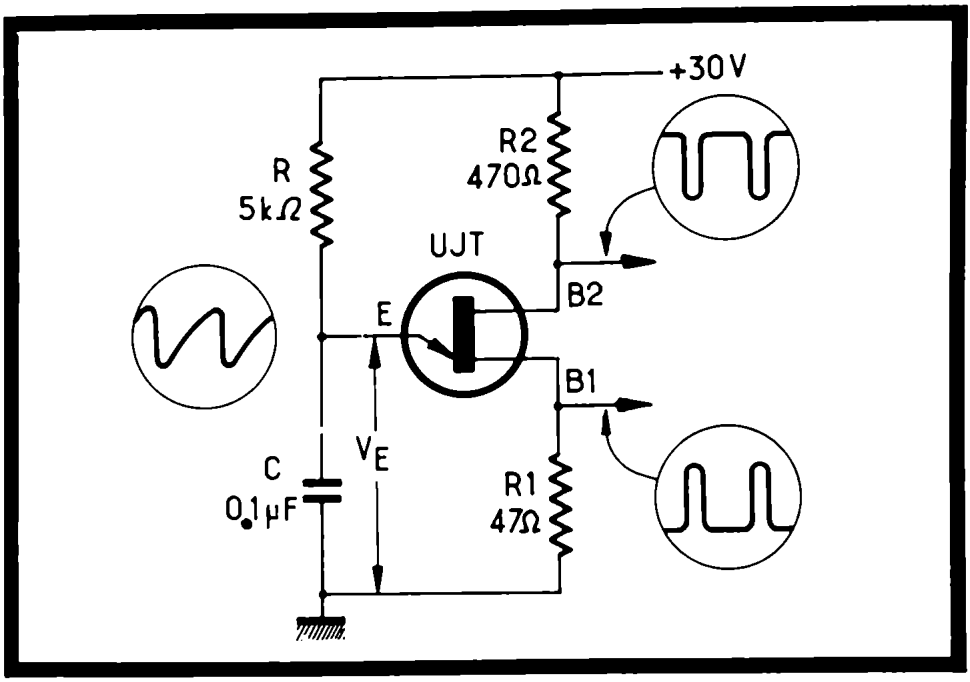


Fig. 3-80. UJT relaxation oscillator circuit.

conduction until the end of the half wave. Thus, RC sets the firing angle of the SCR and the power delivered to the load.

Slightly modified and DC-fed, this circuit becomes a **timer** featuring fair accuracy. (Fig. 3-82.) Pressing normally-closed push-button K removes the short across capacitor C and allows C to start charging. After a time delay, depending

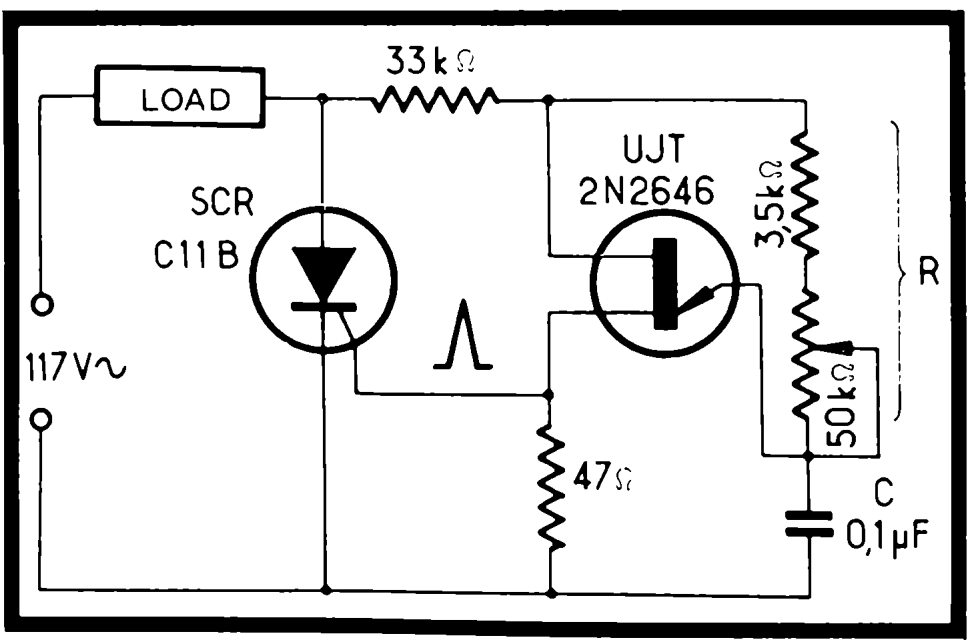


Fig. 3-81. Thyristor control using a unijunction transistor.

on time constant RC, the UJT fires and triggers the SCR. All of the supply voltage (except for the low voltage drop across the SCR) is then applied to the load. Diode D and series resistor R3 deliver a holding current (IO) to the SCR to make sure it stays in the conducting state should the load be disconnected. The delay-determining charging voltage is stabilized by zener diode DZ.

NONLINEAR DEVICES

A conventional ("linear") resistor is defined by Ohm's law as a device in which R equals E over I. This means that should E or I vary, the E over I ratio will always remain constant and equal R. A device is called "nonlinear" if its resistance varies in terms of applied voltage, current drain or any other parameter. (Photoconductors in which the resistance varies with illumination can be considered nonlinear resistors, too.) These devices are extremely useful in regulation and control systems, either for voltage or current stabilization, or rather to exaggerate their variation in order to provide a healthy-amplitude error signal for control purposes. (Only nonlinear resistors are discussed here; nonlinear inductors are described in Chapter 8.)

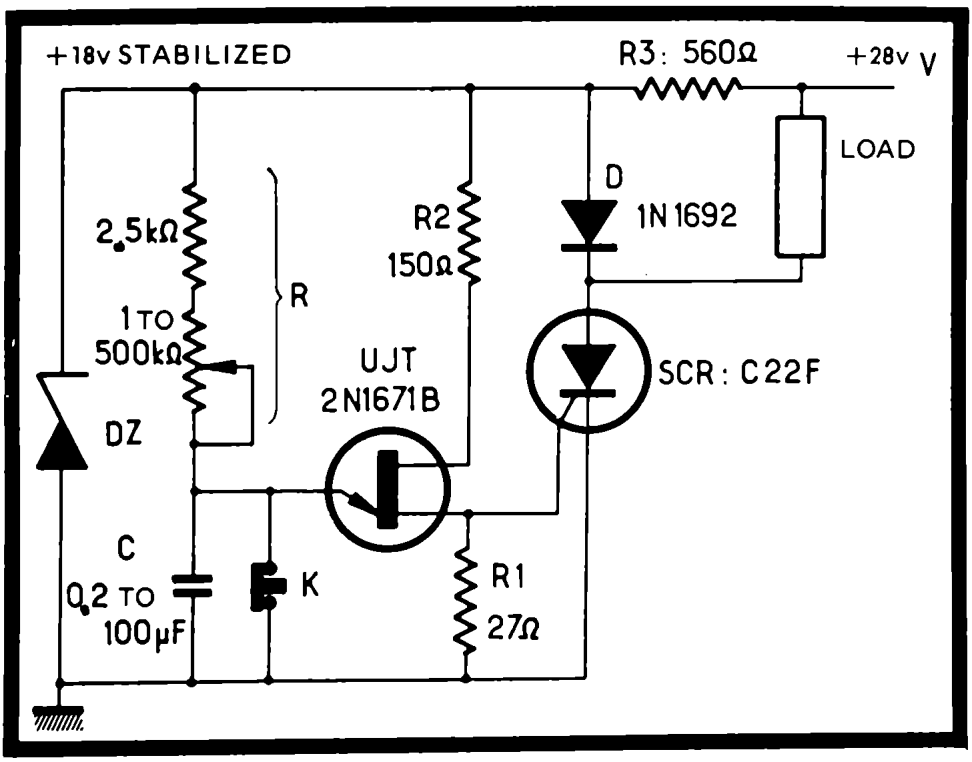


Fig. 3-82. Timer circuit using a UJT-controlled thyristor.

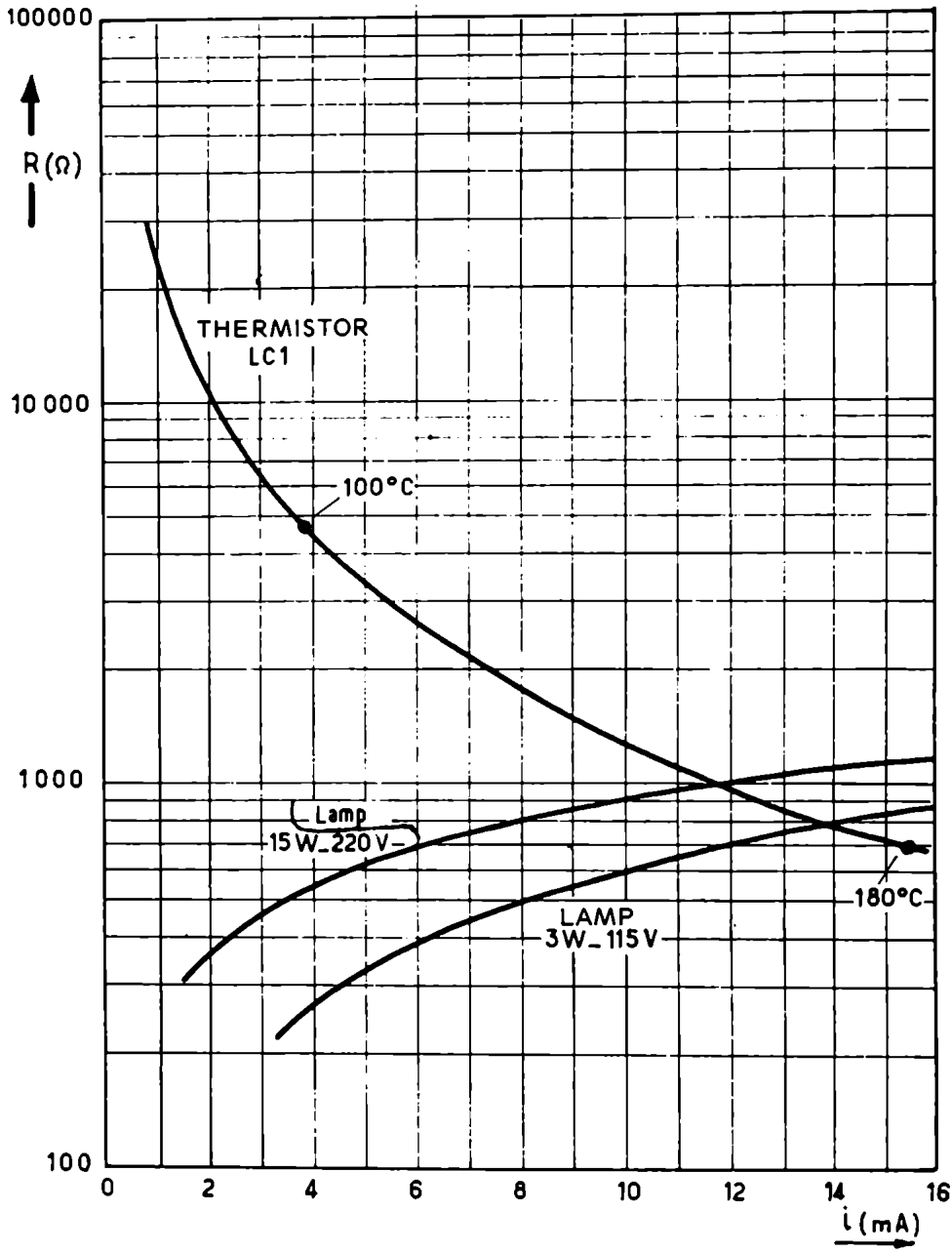


Fig. 3-83. Resistance variation in terms of current for a thermistor and two small-power lamps.

A cheap and easy to use positive-t.c. component is the ordinary metallic-filament lamp. The R over I characteristic curves plotted in Fig. 3-83 represent two lamp types. The resistance of such a lamp varies from about 200 ohms at a cold start to some 1,500 ohms when hot. There is no point in heating the filament to the normal operating temperature, for the

curved segment at the very beginning where nonlinearity is highest is preferred. Low-voltage and high-power lamps present high thermal inertia (and thus slow action) because of the mass of the comparatively thick filament.

Thermistors are nonlinear resistors composed of sintered metallic oxides; they feature a negative comparatively high t.c. of about 4 percent per degree C according to composition. The characteristic of a typical 100,000-ohm thermistor (the rated value is for a 20-degree C or a 68 degree F ambient) is plotted on the same R over I diagram in Fig. 3-83. (Notice the logarithmic ordinate.) This device yields a resistance variation exceeding 100:1, a "honey" for control purposes. Thermistors are available with rated resistances of from a couple of ohms to about 0.2 meg, and they are marketed as discs, rods and tiny pinhead-size pearls for low time constants (some seconds). This component can be heated by a current (Joule effect) as well as by heat radiation or conduction (temperature transducers). There are indirectly heated thermistors composed of the sintered pearl surrounded by a tiny insulated heater winding. Positive-t.c. thermistors are available, too.

The varistor or Thyrite (General Electric) is another nonlinear component; its resistance rapidly decreases when the applied voltage increases. This property makes it useful for overvoltage protection, and Thyrites, made of sintered oxides, make economic high-power lightning-arrestors. Low-power varistors are inverse-parallel connected diodes. Both devices have similar characteristics, including two forward conduction segments in the first and third quadrants; one diode conducts for each polarity.

The nonlinear devices examined above are not polarized; therefore, they are serviceable with AC and DC. Semiconductor, vacuum and gas diodes are highly nonlinear devices; they pass current in the forward direction, but very poorly or even not at all in the reverse (blocking) direction.

On the blocking side, the vacuum tube is best, for there is no possibility for a reverse current. Silicon is nearly as good, especially if the voltage applied is low compared to the rated peak inverse voltage. Germanium has an slightly higher reverse current, and selenium has appreciably more. However, all types are satisfactory if they are chosen in terms of the rectifying characteristics required.

The zener diode is a silicon diode featuring a particular reverse characteristic as shown in Fig. 3-84. (It is this branch that is used for voltage regulation. Notice the forward conduction is the same as that of any other silicon diode.) For low voltages, the reverse current practically is zero. Then sud-

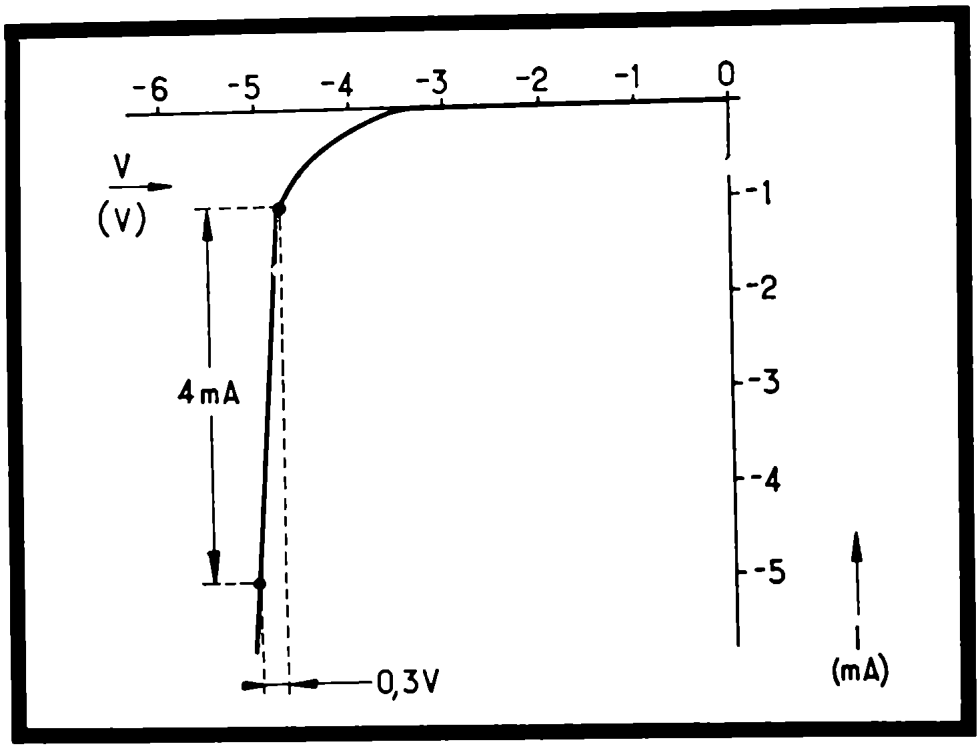


Fig. 3-84. Curve showing the steep reverse characteristic of a zener diode.

denly, for a voltage exceeding a threshold, called **zener voltage** (about -5v on the diagram), the current increases rapidly (a voltage increase of 0.3v corresponds to a current increase of as much as 4 ma, making for a differential resistance of dV over dI equals 75 ohms, and sufficient series resistance must be provided to limit the current to a safe value.

With a series resistor Rs, the Zener diode (DZ) forms a voltage stabilizer (Fig. 3-85). This actually is an attenuator coil. If Rs is 7,500 ohms, with a differential resistance of 75 ohms for DZ, input voltage fluctuations will be reduced to one one-hundredth of the original amplitude. Permuting Rs and DZ results in a circuit that exaggerates the input fluctuations and thus yields a fair-amplitude error signal (Fig. 3-86). The resistor is of low value here in regard to the differential resistance. Thus, a constant voltage drop, preferably only slightly lower than the input voltage, develops across DZ, leaving mainly the fluctuating component across the output. For control purposes, it is necessary only to amplify this component and reinject it into the input.

Zener diodes are made for stabilizing voltages up to 100, but devices designed for zener voltages of 5 to 6 feature the

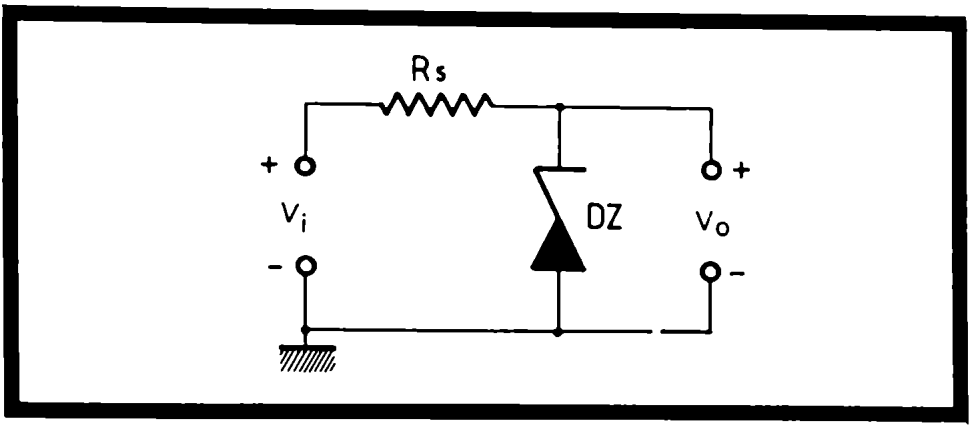


Fig. 3-85. Zener diode used as a voltage stabilizer.

lowest differential resistance and a practically zero temperature coefficient and are, therefore, preferred. For voltages between 65 and 150, gas regulator tubes (glow lamps) are used. Nowadays, Zener diodes have superseded glow lamps as regulators in most applications, but glow-type indicators, ofcourse, will survive as long as gallium arsenide injection diodes are not mass-produced.

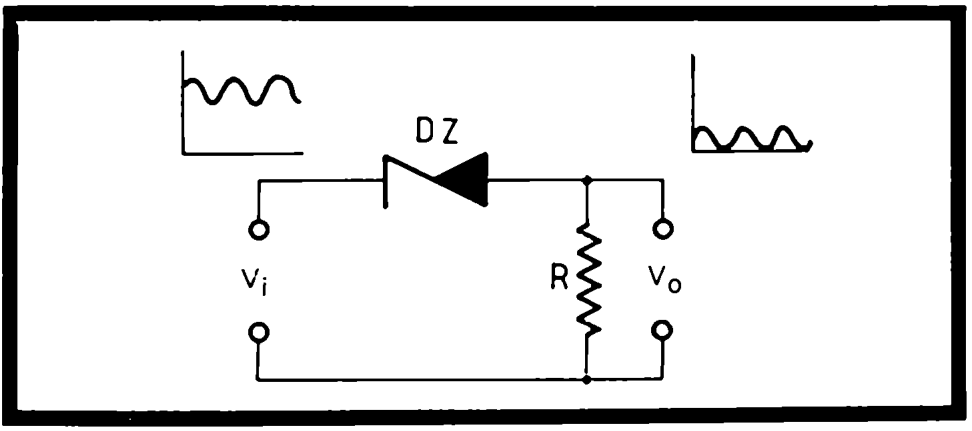


Fig. 3-86. This zener circuit exaggerates voltage fluctuations and can be used to provide a convenient error signal.

CHAPTER 4

Automatic Inspection, Sorting & Counting

Modern manufacturing operations are accompanied by a host of accessory—though vital—functions such as quality-control inspection, sorting to reject inadequate parts and counting finished products. In delicate and sophisticated product lines, such as semiconductor manufacturing, about every manufacturing step may be followed by an inspection to avoid spending time down the line on defective items.

Automatic counting has come of age. Nowadays, for a person to do nothing all day long but count items carried by a conveyor belt cannot even be imagined. An automatic counter never gets tired; it can count at any speed without making mistakes and obviously is more economical than its human counterpart. A suitable (generally photoelectric) transducer is required to detect the items to be counted and to issue counting pulses accordingly. The resulting count may be displayed and-or recorded on magnetic or paper tape and-or printed out. Integrating the results of all counting stations immediately makes available a complete daily production record.

Automatic counting is straight-forward, but automatic inspection is a more or less complicated function. The number of things that can be inspected, the inspection modes and the rules governing the accept-reject decision are infinite. Automatic inspection excels in simple routine such operations as dimensional monitoring. Gauging fabricated items in a production line is straight forward, accurate and economical. On the other hand, consider the feasibility of a machine for the final inspection of a piece of apparel! A trained worker probably could do the job more economically. An automatic apparel inspection equipment certainly can be realized, but at what price? It could even decide if the finished job looks stylish!

A sorting device is a system that takes action to carry out the accept-reject decision of an associated inspection device. This action actually consists of separating rejected from accepted batches by directing them to appropriate bins, piles or conveyor belts. These devices, analogous to railway

switches, are electrically, hydraulically or pneumatically operated actuators.

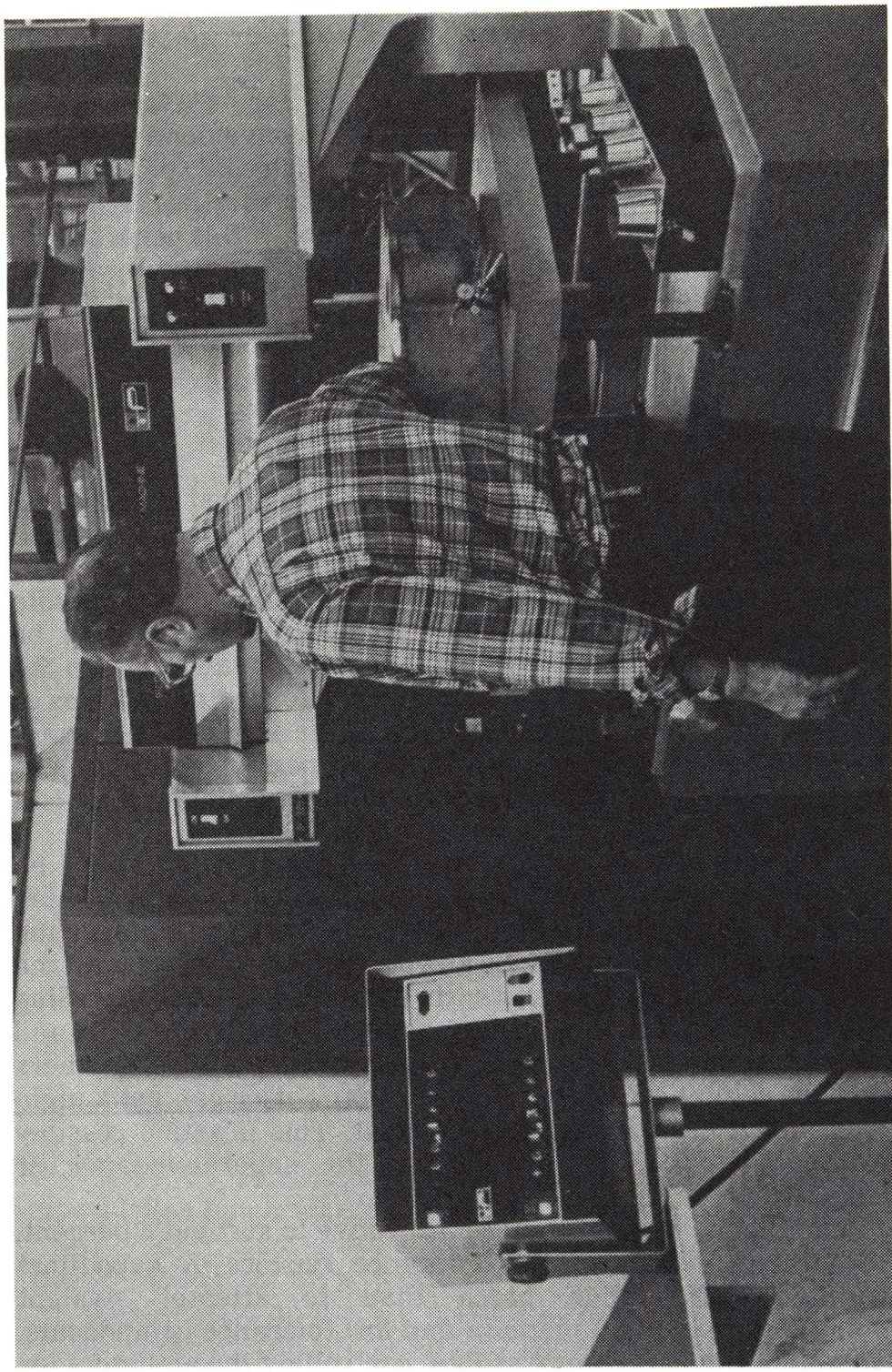
Besides mere product checking for conformity to specifications, the inspection device is conveniently used to check for correct operation of the up-stream part of the production line. Repetitive dimensional errors, for instance, may be due to a broken or worn-down tool. If this occurs at an unusual rate, the machine not only rejects the faulty items, but also shuts down the line and triggers an alarm. Thus, automatic inspection is a first step to automatic production supervision.

STRUCTURE OF INSPECTION SYSTEMS

To avoid any ambiguity in interpreting test results, the inspection system is designed to yield definite decisions for a given set of conditions, such as the simple (binary) **accept-reject** or **go-no go** indications, leading to the simplest possible 2-way sorting actuator control. In other cases, it may be desirable to know if the rejected items are on the low or on the high side. In the case of machined parts, for instance, items of insufficient length or diameter must be scrapped; it may, however, be economical to reprocess parts showing a size excess. The corresponding inspection and sorting system is a 3-way system: **high, go** and **low** (Fig. 4-1).

Two transistors of the same type, though both may be acceptable, may offer somewhat different current gain (beta) values. Manufacturers have not succeeded in reducing the beta spread in production. So an automatic system tests all the devices leaving the line and sorts them into batches according to beta values, ranging, for instance, between 20 and 40, 41 and 60, and 61 and 80. These batches get different type numbers and are sold at different prices, so that a user can get a high-beta unit if he is willing to pay for it. This practice possibly explains one of the reasons for the superabundance of transistor type numbers.

Similarly, it may be more attractive to sort out 5 percent, 10 percent and 20 percent tolerance reistors or capacitors within a production batch than to attempt straight manufacture of high, medium and low-tolerance components. An automatic percentage bridge, controlling the actuators for sorting and color-code painting, will do the trick. Such an automatic inspection system could include a luminous display flashing the batching decision, and counters totaling the items of each batch and optically displaying the counts. Preset counters also can signal completion of the required number of items in a batch.



At the diesel engine manufacturing plant of Cummins Engine Co., Columbus, Indiana, a coordinate measuring machine is used to check the accuracy of borings on a flywheel housing. The device is used primarily as a final audit on parts already inspected by on-line personnel. Developed by Potter Instrument Co.

There are inspection systems of increasing complexity for checking only one parameter or many magnitudes. Using appropriate transducers, chemical, physical, mechanical or other variables are converted into an electric current, voltage or resistance, so that the "works" of inspection systems for all kinds of industries are consistently electronic. Generally speaking, only the parameters vital to correct device operation are tested, and there is no point in exaggerating measurement accuracy, for the system's reliability decreases while its cost increases. There are test fixtures and benches for single components, subassemblies such as oscillators and amplifiers, chassis and plug-in drawers, and even complete systems. If successive testing operations are to be performed, the inspection system is to be **programmed**; i.e., instructed to proceed step by step according to directions. A basic program is required even if only one parameter is to be tested.

STEPPING SWITCHES FOR PROGRAMMED INSPECTION

A stepping switch includes several decks of stationary contacts (D, Fig. 4-2) swept by concentrically-mounted rotating contact arms or wipers (W). Gathering pawl P, driven by armature A of electromagnet M, engages with toothed wheel T, which is integral with the wiper spindle. Each time the magnet is de-energized, it releases its armature; spring S pulls down the pawl, and the wipers proceed to the following contact. Thus, energizing M sets the mechanism for a new step carried out upon armature release; thus, the pawl pull-down force is independent of the energizing voltage.

Stepping (or uniselector) switches are very flexible, and many useful types of operation can be obtained by skillful

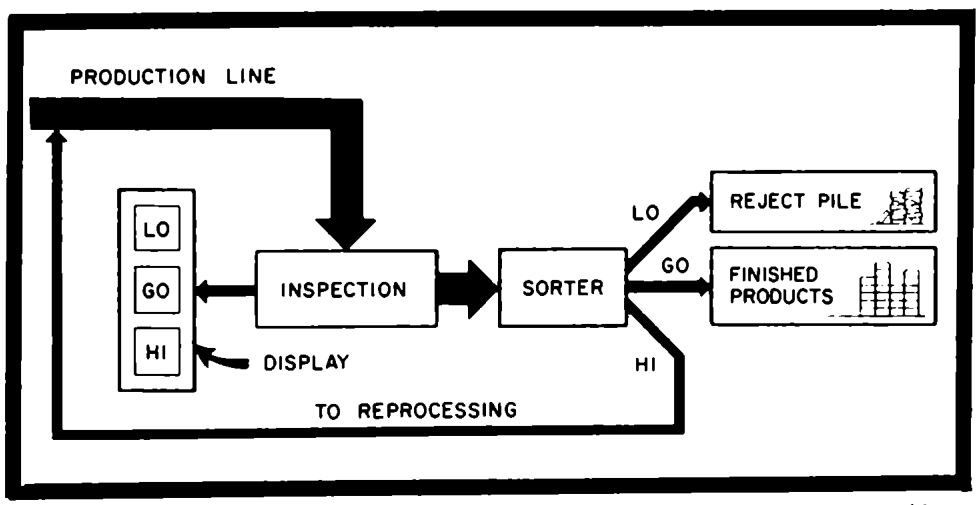


Fig. 4-1. Operating diagram of an automatic inspection and sorting system.

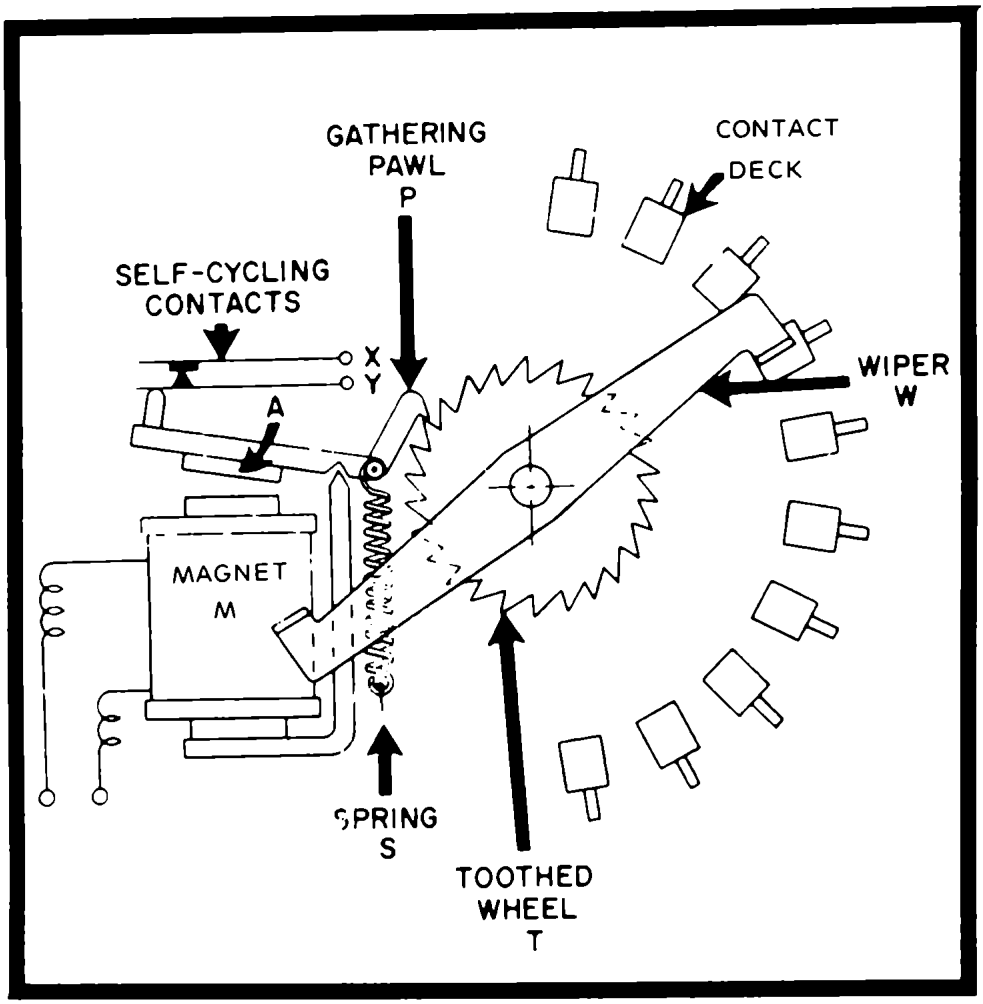


Fig. 4-2. Drawing of a magnet-driven rotary selector switch.

circuit design. The "stepper" in Fig. 4-2 carries two normally closed contacts, X and Y, which are opened each time the coil is energized. This allows for self-cycling (Fig. 4-3). If coil M, contacts X and Y and the power supply are series-connected, the operating current is interrupted as soon as armature A is attracted by M. This de-energizes M, X and Y close again, and, provided the circuit remains closed, the wipers step around again and again. The dwell time of the wiper on the contacts can be lengthened by shunting a capacitor (C) with sufficient capacitance (say, 100 mfd) across M; the capacitor causes a slow current buildup. Using several capacitors and an additional contact deck, it is even possible to obtain different dwell times for the various contacts.

Homing, or coming to rest at a given position, is accomplished by tying together contacts K1, K2, K3, etc., of one deck and connecting them in series with the driving circuits as

shown in Fig. 4-3. Notice that K5 is not connected to the other contacts. Thus, the wiper will stop at K5, the "homing contact"; therefore, K5 is the starting point of every operating cycle. A new cycle can be started by momentarily closing pushbutton switch PB. In an automated inspection system, the start circuit can be closed mechanically with a switch incorporated in the jig or holding fixture of the item to be tested and closed as soon as the device is positioned for testing.

For quick repair of defective subassemblies, the faulty circuit is identified by the number of the corresponding test. To stop the stepper at the corresponding contact, the checking unit opens a normally closed contact for a no-go decision. Inserting this contact between the corresponding K contact and the common link will stop rotation at the faulty test position and display its number. The operator can note the number on a tag accompanying the subassembly, and press a button to resume the homing operation. If the production line is highly automated, speedy clearing-up upon fault detection can be carried out by simply printing the check number on the identification tag without stoppage of the inspection system.

The solid-state circuit in Fig. 4-4 allows for varying within wide limits (from one step in 5 seconds to 15 steps per second) the stepper progressing speed. In this AC-fed circuit, the stepper coil is energized by a UJT-triggered thyristor (SCR). The UJT firing delay varies with time constant RC, which is adjustable by means of variable resistor R. When the supply connections are closed, contact XY is closed, but the SCR is not conducting and coil M is not energized. After a time delay

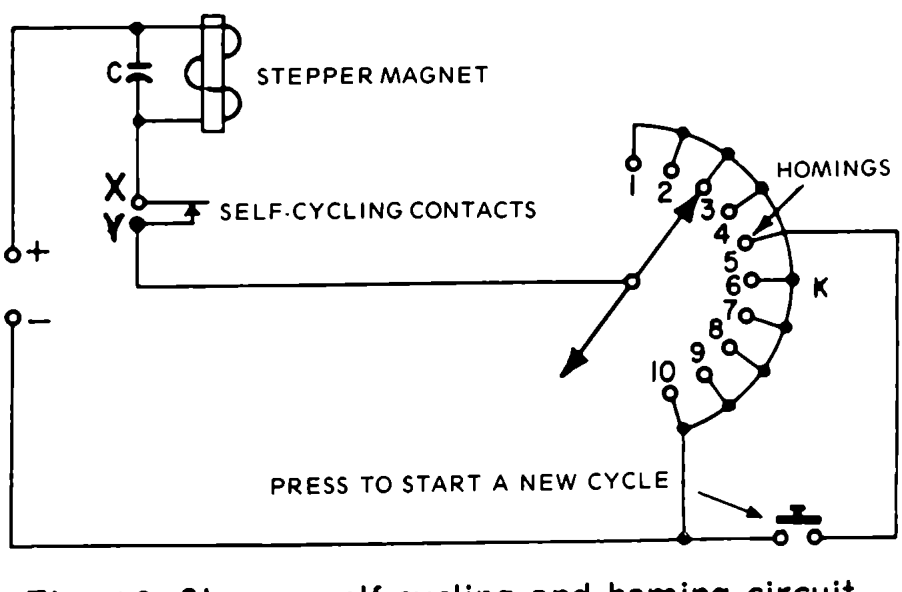


Fig. 4-3. Stepper self-cycling and homing circuit.

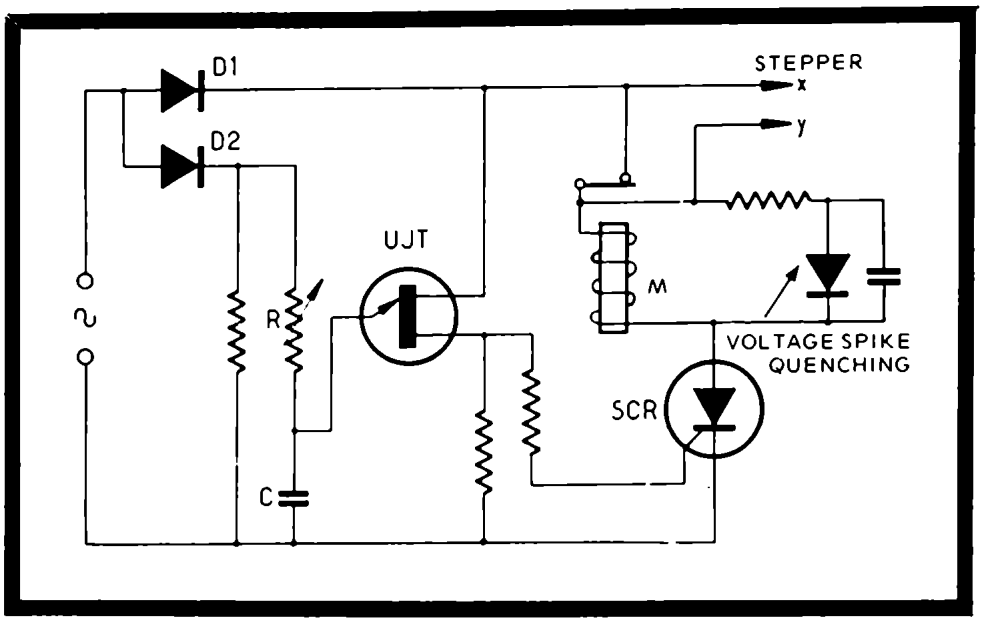


Fig. 4-4. Schematic of a thyristor-operated stepper device.

t, the UJT fires and issues a triggering pulse to the SCR. The thyristor strikes and energizes M, but this interrupts XY and, thereby, de-energizes M and, simultaneously, the SCR. The wiper then proceeds one step forward. Resistor R is external to the control subassembly; therefore, it is possible, by means of an additional contact deck, to connect an individual resistance to every contact position and thus to set at will the dwell time for every test position. A **programmed timer** results, which assigns the proper duration to each of a series of consecutive actions.

AUTOMATIC INSPECTION OF MULTICONDUCTOR CABLES

Electric appliances and electronic equipment leaving the assembly line must be checked for **continuity** of wiring as well as for shorts between conductors and to ground. This test can be time-consuming if many conductors are involved as in wiring harnesses. The connector-terminated multiconductor cable is a particular problem.

A set for automatic 6-conductor cable testing by means of a homing stepping switch is diagrammed in Fig. 4-5. By means of contact decks a to g, every pin of one connector is successively tested for continuity with the corresponding pin of the other connector, plus a short circuit check between each pin and the five remaining pins tied together. Thyristors SCR1 (continuity) and SCR2 (shorts) are connected according to A

and B in Fig. 3-73; they are normally de-energized. If the connection between the corresponding pins of both connectors is interrupted, the SCR1 gate would no longer be grounded, firing the thyristor and energizing the relay. The relay contacts would interrupt the stepper feed, thus stopping the sweeper at the fault position.

If a short or an important leak develops between the selected pin and all others tied together, SCR2 would fire and energize the relay, stopping the stepper at the corresponding

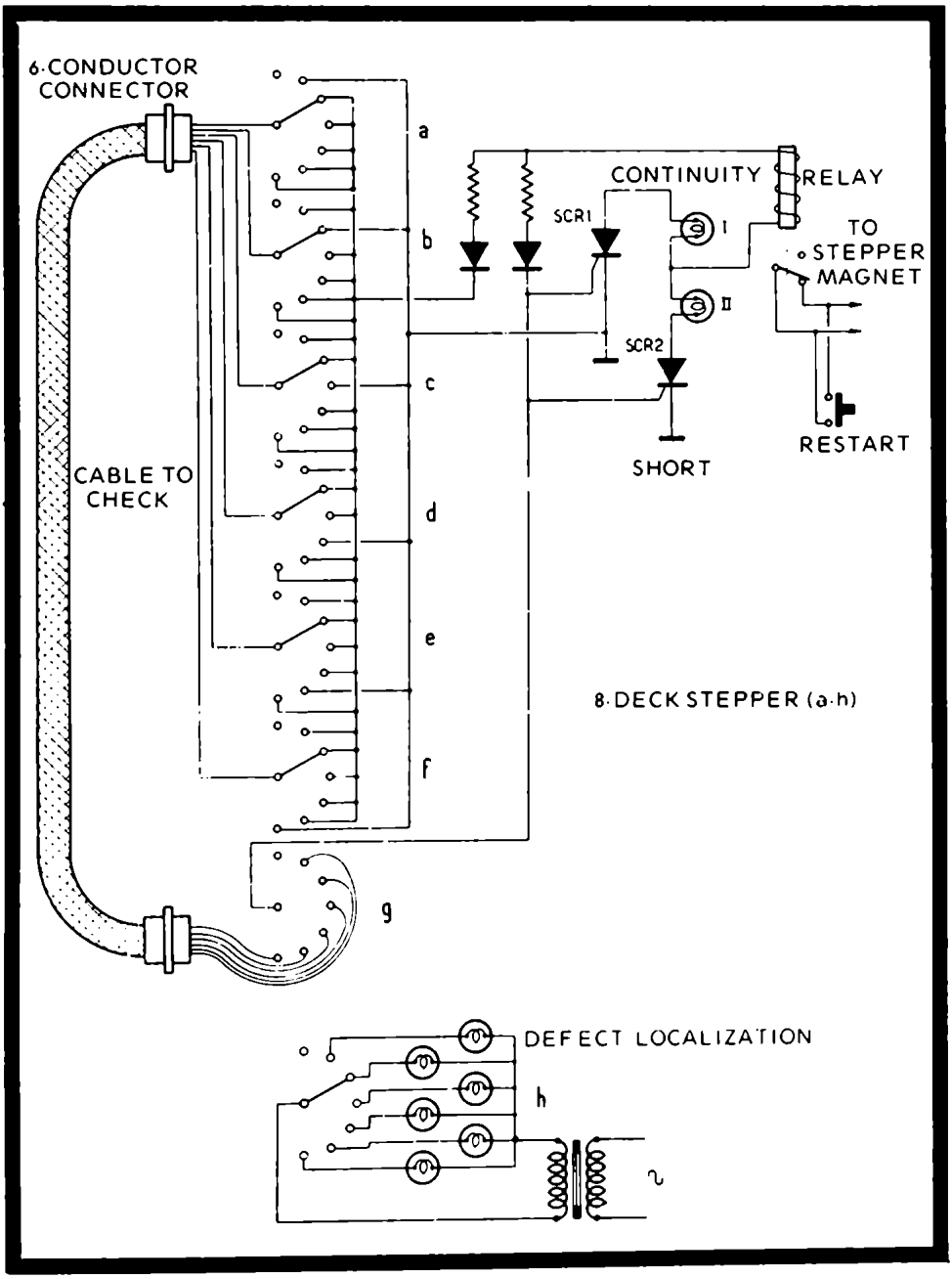


Fig. 4-5. Automatic unit designed to check multiconductor cables.

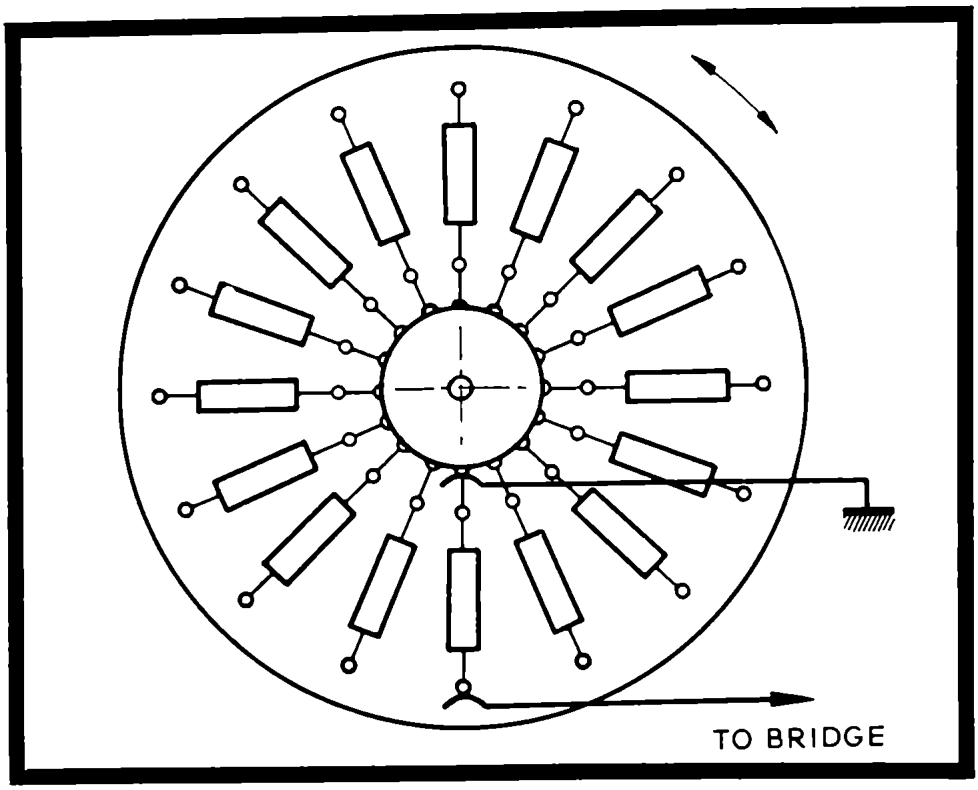


Fig. 4-6. Turntable used to insert resistors in an automatic tester.

position. Control lamps I and II, series-connected with SCR1 and SCR2, tell the nature of the fault—discontinuity or short. The numbered control lamps, connected by contact deck h, identify the faulty position. If everything is OK, the stepper proceeds to its homing point and no light remains on.

AUTOMATIC TESTING OF PASSIVE COMPONENTS

Components are tested by the manufacturer at the end of the production line and sometimes again by the equipment manufacturer. Since the inspecting equipment has to pass a large number of components (resistors, for instance) with little manpower, if any, to be economically feasible, a large part of the testing machine is made up of devices designed for automatic loading. This is usually done by means of a hopper feeding the resistors in line by means of a vibrating funnel with an inner spiral groove. Once the testing is accomplished, the component is directed to the selected chute by some type of actuator. If no automatic loader is available, the resistors can be hand-fed to the holding jigs placed radially on a rotating disc (Fig. 4-6). An elastic sliding contact connects to all components successively as the disc is rotated.

Unknown resistor Rx is connected in a Wheatstone bridge made up of equal resistors R1 and R2 (Fig. 4-7), forming proportional arms a over b equals 1, and standard resistor R3 equaling the rated value of Rx. A symmetrical DC amplifier is connected to one diagonal of the bridge and its output feeds a biased relay. If Rx is within the specified tolerance, opposing flux developed by the collector currents of transistors Q5 and Q6 balance, the relay is not energized, and no contact closes; accordingly, the component is deposited in the "go" bin. If Rx tests high, there is an unbalance which closes contact K1 to energize electromechanical actuator. A1, directing the component to the "high" bin; similarly, a low testing component closes contact K2 to energize actuator A2 which switches the item to the "low" bin. The amplifier sorting tolerance is set by variable resistor R6.

Reactive components, such as inductors and capacitors, are tested on some type of AC bridge fed by a low-frequency source (from 50 to 4,000 Hz) as outlined in Chapter 1. Remember, there are two balance conditions to be satisfied—a reactive one and a resistive one, but as the magnitude variations encountered in industrial control are generally small, balancing only one component may do. An automatic 2-variable balancing control is, of course, more complicated to set up.

An unknown capacitor, Cx, in series with a resistive component, Rx, can be compared to a standard capacitor, Cs, in series with a variable resistor, Rs, for loss compensation on the Wien bridge in Fig. 4-8. Balance is obtained by ratio arms R1R2 if Cx over Cs equals Rs over Rx equals R1 over R2. Straight rectification of the amplified AC error signal would yield a DC current that increases with unbalance. Of course, the unbalance direction information would be lost. The phase discriminator in Fig. 4-8 is fed a reference voltage by the center-tapped secondary of the bridge supply transformer, and the center-zero galvanometer G deflection indicates the magnitude as well as the direction of the unbalance. If this information speeds up manual bridge balancing, it is essential for instructing automatic balancing devices that are to be told in what direction to proceed.

The **bridge modulator** in Fig. 3-31 is reversible and serviceable as a discriminator, too. The same holds for the **ring modulator** with four diodes connected the same way around (Fig. 4-9). According to the reference voltage half cycle, diodes V1 and V2 or V3 and V4 conduct, connecting input Terminal P to Point M or N of voltage divider R1R2. The divider midpoint Q remains tied to the other input terminal.

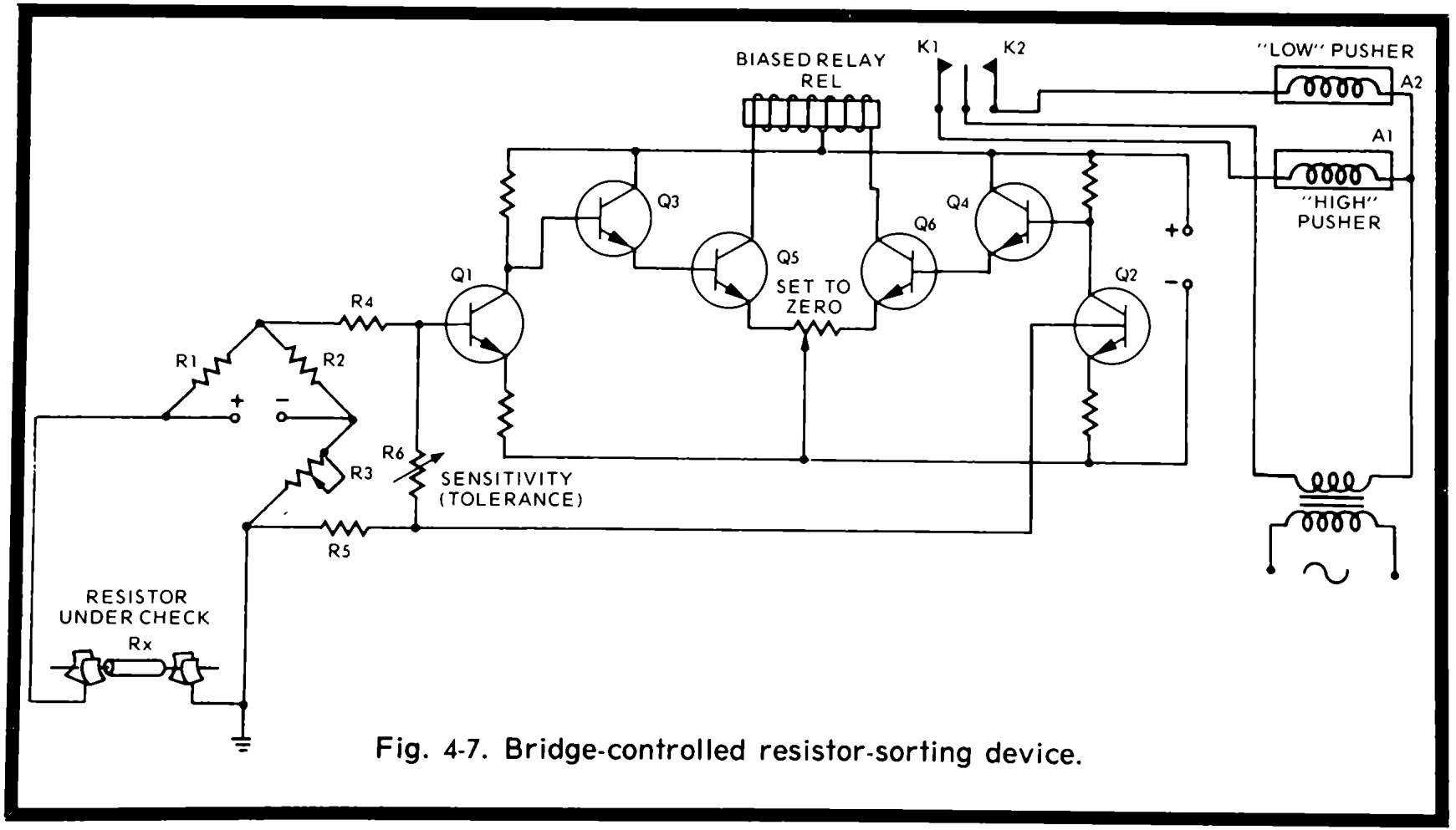


Fig. 4-7. Bridge-controlled resistor-sorting device.

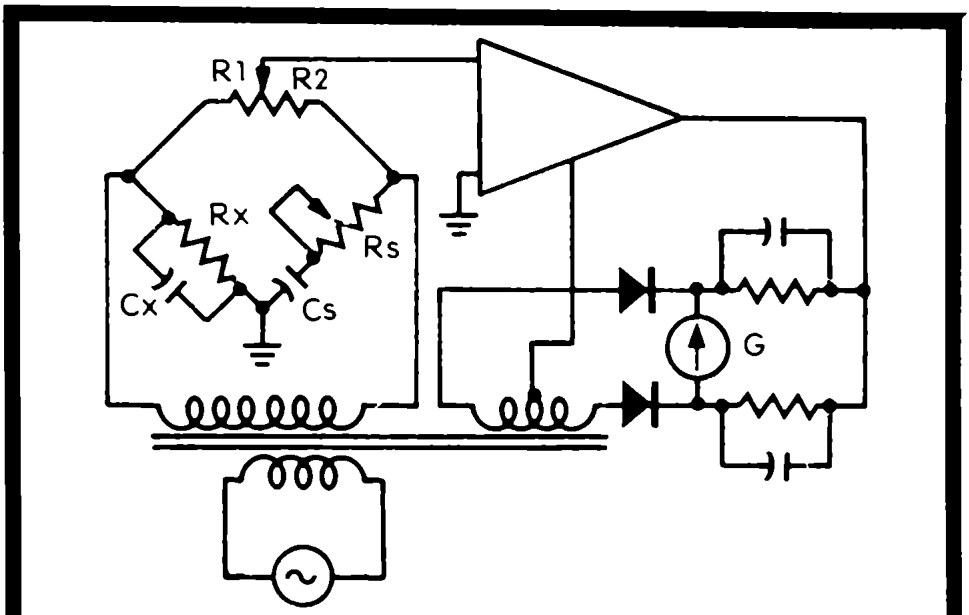


Fig. 4-8. Wien bridge capacitor checker.

This results in a pulsed unipolar output where the polarity depends on the error voltage phase or on the direction of unbalance. The circuit behaves like a double-pole double-throw switch, reversing the signal polarity at each reference voltage half cycle.

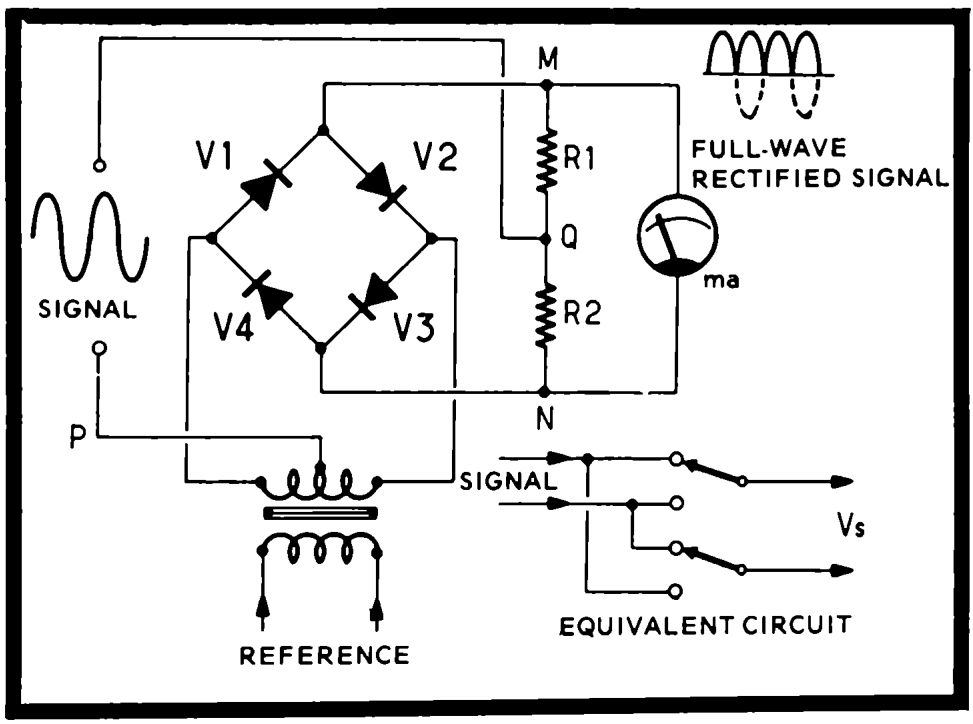


Fig. 4-9. Ring modulator circuit used as a phase discriminator for null detectors.

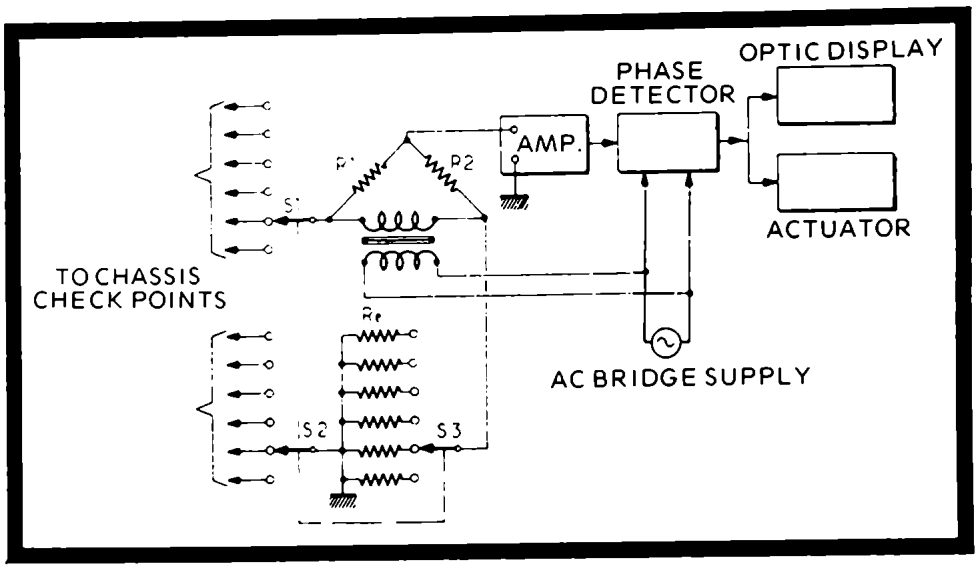


Fig. 4-10. Bridge-controlled automatic subassembly tester.

TESTING CHASSIS AND SUBASSEMBLIES

Automatic systems for testing chassis and subassemblies can be designed by putting together a stepping switch, a self-balancing bridge, signaling lights and some electrically controlled actuator such as a pusher (Fig. 4-10). Multiconductor cables tie the bridge to selected test points by means of suitable connectors. Contact levels S1 and S2 of the stepper successively connect the chassis resistors or impedances into the bridge, while level S3 inserts the corresponding standards. Only resistors are shown, but capacitors, inductors and combinations thereof can be used as well. The automatic cycling and homing system (not shown) is the same as described above.

This operation is a static test of the assembled components (with no supply voltage applied) to detect shorts, open circuits and wrong-value components; this is a necessity, especially if the components have not been tested before assembly. If these tests are all right (and only then), the gear is tested in operating condition, with normal supply voltages applied; this is a dynamic testwhich consists of measuring DC and signal voltages and currents at transistor electrodes. Digital meters are preferred if the test bench is run by an operator, for they provide speedy, error-free readings. Of course dynamic tests can be automated, too, but because of the complicated test equipment involved, a human operator able of discerning qualitative factors may be favored.

The stepper introduces switching transients in the amplifier input, resulting in erroneous operation of actuating and display devices. Slow-acting relays may be an economic solution to this problem, for they will delay switching sufficiently for the transient to die away. Such relays comprise a copper slug fixed on the iron core in a way to delay pull-in or dropout by inducing eddy currents in the copper mass. In static switching circuits, delay is obtained even more easily by a simple RC circuit in series with the control electrode, gate or base.

SELF-BALANCING BRIDGES AND POTENTIOMETERS

In a test operation on a production line, manual bridge adjustment for each measurement is not economically feasible. It is not difficult, however, to use the amplified error signal to control a small motor to drive the measurement potentiometer. Such a self-balancing Wheatstone bridge is illustrated by the diagram in Fig. 4-11. The present bridge is AC-fed, its balancing potentiometer R is driven by a geared-down 2-phase induction motor. Such a motor is reversible, and its direction of rotation depends on the phase difference between reference winding L1 and control winding L2.

L1 is fed by the AC source used to power the bridge (or alternately by the chopper reference voltage, in the case of a DC-powered bridge). By a capacitor (C) of adequate value, the current energizing L1 is in quadrature in relationship to the reference voltage. Because the purely resistive bridge introduces no phase shift, the amplified error voltage developed across control winding L2 is in quadrature in respect to the current flowing through L1, leading or lagging according to the direction of bridge unbalance. This makes the motor rotate in one direction for an unbalance on the low side and in the opposite sense for "high." As soon as the equipment is energized, the motor starts running, setting the potentiometer to balance the bridge, and stops upon reaching this condition. (Just like in other servosystems, there are transient and stability problems to be solved, and the system must be critically (or slightly undercritically) damped as indicated in the first chapter).

The small motor is powerful enough to drive, in addition to the potentiometer, a pointer moving before a large scale for convenient distant reading, and possibly also the pen of a graphic recorder which traces the variation in terms of elapsed time. Such a device goes far beyond mere component testing. The resistor designated by Rx can stand for a resistive transducer that translates the variations of an industrial

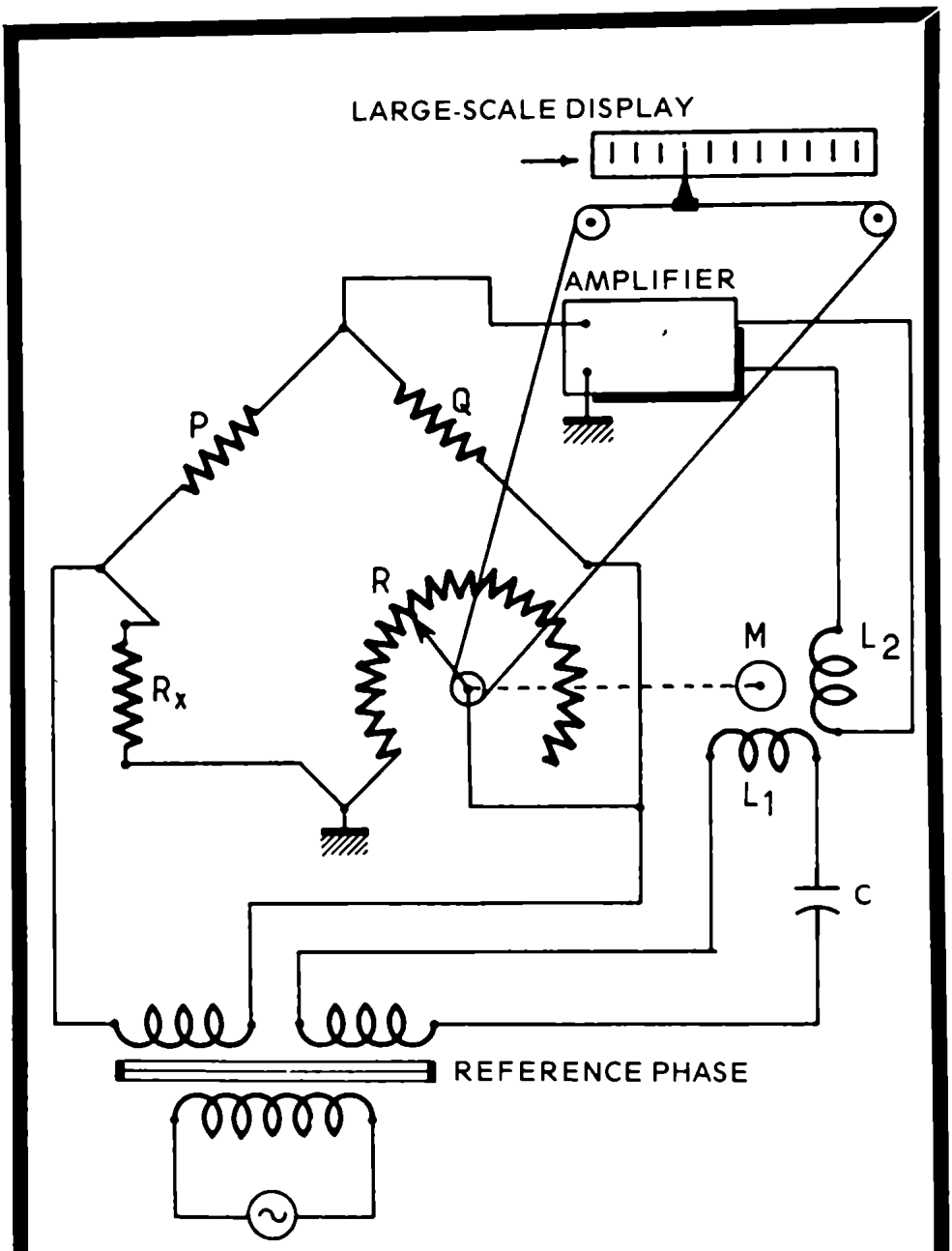


Fig. 4-11. Self-balancing Wheatstone bridge circuit.

process parameter: temperature, pressure, volume, flow or other. This variation is immediately displayed on the scale and recorded on the paper strip, so that a process record is available without delay. Sufficient power provided, the motor can even be made to act upon a control agent in order to restore a specified condition. So we have **programmed process control or automation.**

The potentiometer method of voltage measurement (the other null method) similarly leads to self-balancing (Fig. 4-

12). Voltage V to be measured is opposed to voltage drop e sampled off potentiometer P, whose slider is driven by a geared-down 2-phase motor. The voltage difference V - e is applied to a chopper modulator, and the chopped wave obtained is amplified to energize control winding L2. Reference winding L1 is fed the reference voltage through quadrature capacitor C. The difference voltage, amplified, causes the motor to rotate in order to drive the potentiometer slider to null the voltage difference V - e, resulting in V equals e. Once the balance condition is achieved, the motor stops. Again, the motor can control a large display, a recorder, and possibly a process control agent.

Both of these devices are purely **analog**, as opposed to digital or discrete devices such as the counters we discuss later in this chapter. While digital methods and systems are making great strides, there is no point in digitalizing systematically without need. It should be a simple matter, however, to link a digital angular position encoder—an analog-

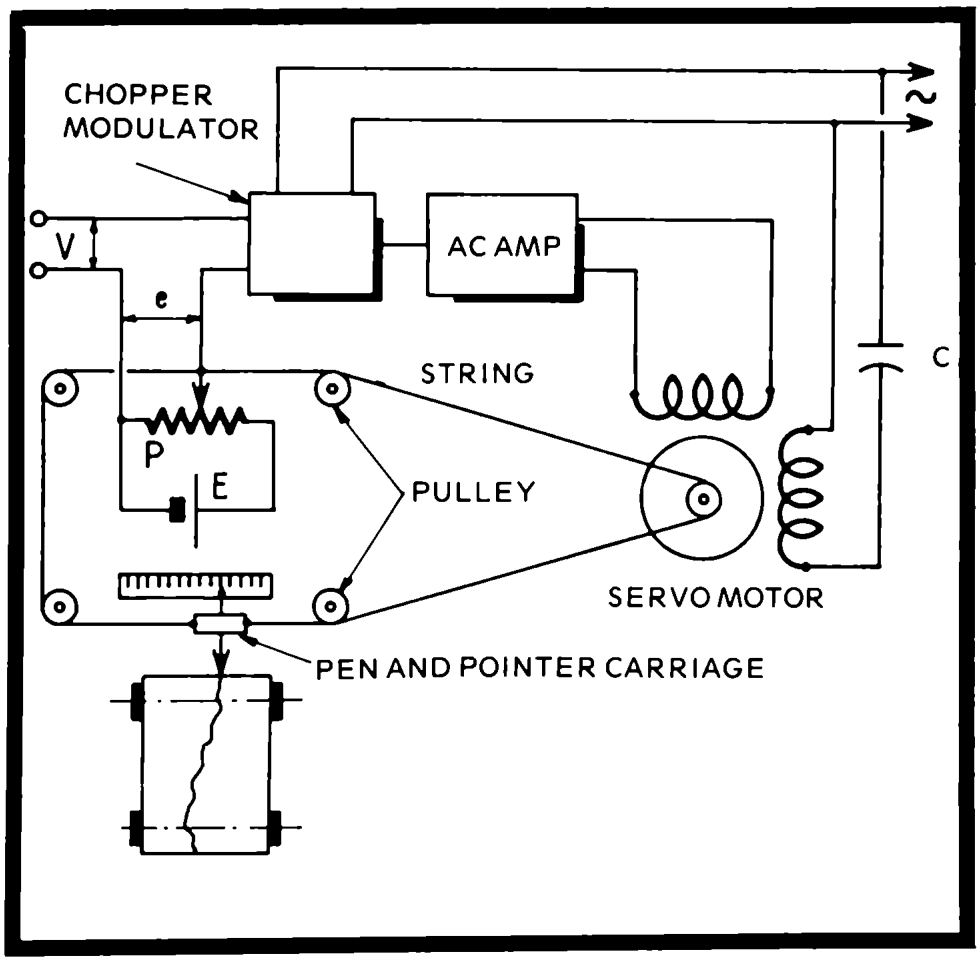


Fig. 4-12. Self-balancing potentiometer designed for voltage comparison.

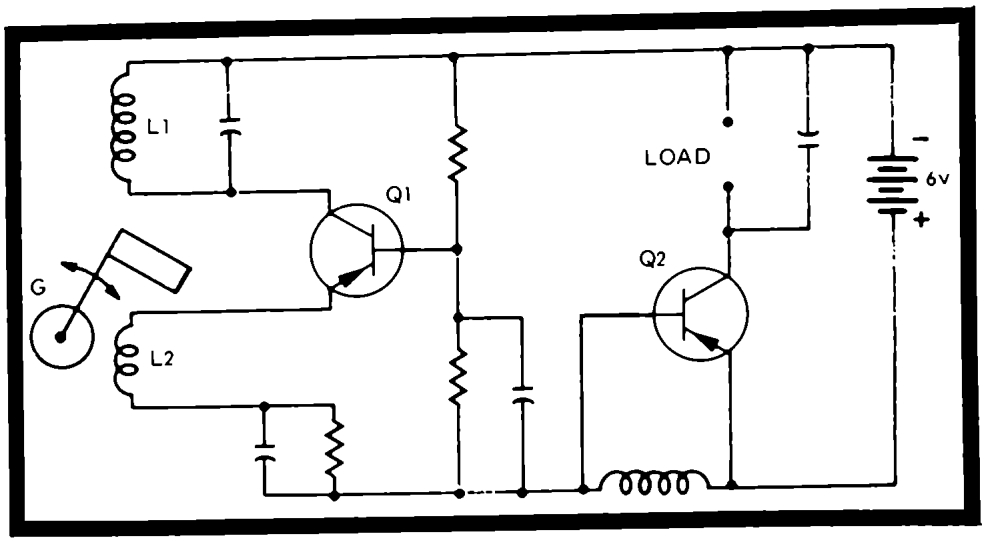


Fig. 4-13. In a galvanometric amplifier, a transistor oscillator is controlled by a galvanometer-driven metal vane.

to-digital converter—to the potentiometer spindle to obtain a digital readout usable for direct numeral display (remote, if required) and for print-out; this not only makes for quick and error-free reading, but the possible accuracy is much higher.

GALVANOMETRIC AMPLIFIERS

The error voltage of a nulling device or the signal of a transducer, generally, is too small to directly operate a sturdy display, a graphic recorder, or even a process parameter control agent. The self-balancers above are one solution to this problem; another is the galvanometric amplifier that can lead to a simpler and less costly solution.

In galvanometric amplifiers, the signal-sensitive means is a taut-band galvanometer with a sensitivity sufficiently high to dispense with an amplifier. A potentiometer is used, but instead of placing the unknown in opposition to a constant current passing in a variable resistor (potentiometer), a compensation current I traversing a fixed resistor R is varied in terms of galvanometer current i, fed by source S, composed of an LC transistor oscillator Q1 and a rectifier-amplifier Q2 (Fig. 4-13). Between windings L1 and L2 of this oscillator (tuned to about 400 kHz) a small metallic vane rotates freely. The vane is solidly connected with the moving coil of d'Arsonval galvanometer G. An increasing i drives the vane between L1 and L2, modifying the inductive coupling of the oscillator coils, thus reducing the oscillator current, which is rectified and amplified to become I.

The operation of this potentiometric instrument as a voltage amplifier is illustrated by Fig. 4-14A. When voltage Vi is applied to the input terminals, G is traversed by a current i equals (Vi - RI) divided by G. This rotates the moving coil of G until Vi equals RI and i equals 0. As soon as compensation i is reached, there is no more input current, meaning an infinite galvanometer sensitivity. Practically, losses by mechanical friction and others cannot be avoided altogether, so that the zero-current condition is only approached.

Use of the instrument as a current amplifier is schematically shown in Fig. 4-14B. G controls I to set both input terminals to the same potential, resulting in RI equals (R + R1) i, or I equals i (R + R1) divided by R. This factor, (R + R1) divided by R, represents the current amplification of the instrument. The voltage drop across the device input, theoretically, is zero; it is very low in practice. The load current thus follows the input voltage variations. Replacing Vi with a source of constant emf will result in a constant-current load supply. A compact transistorized galvanometric amplifier is available (Brion Leroux, France). It looks like an octal-based electrochemical capacitor.

Another type of galvanometric amplifier using tiny photoconductors is illustrated in Fig. 4-15. A taut-band mirror galvanometer is used here. The light beam it reflects forms a

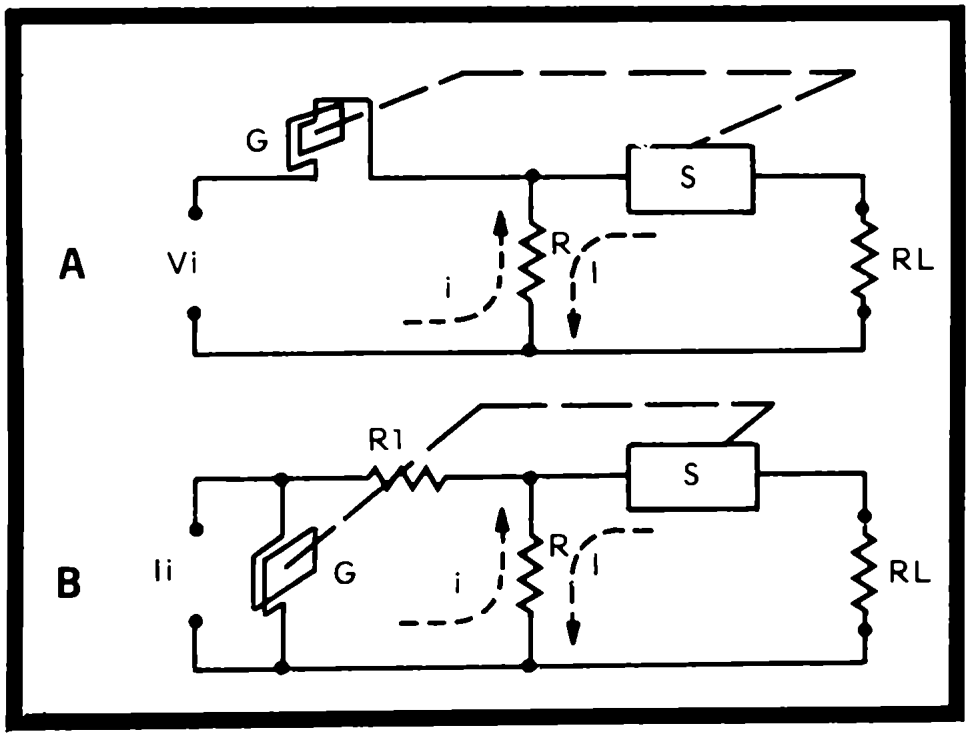


Fig. 4-14. Galvanometric amplifier connected as a voltage (A) and current (B) amplifier.

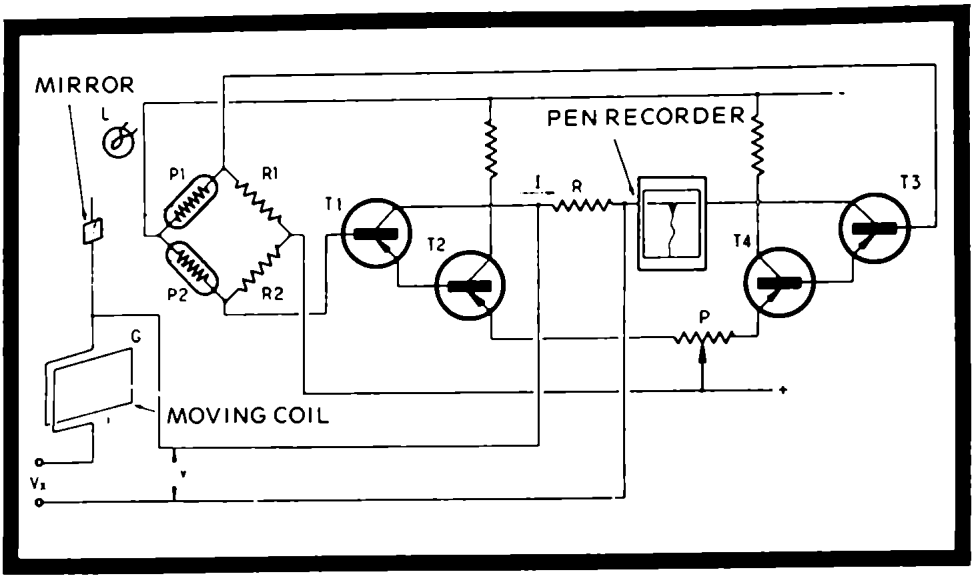


Fig. 4-15. Galvanometric amplifier using a photoconductor bridge.

spot centered between two photoconductors P1 and P2, placed side-by-side and bridge-connected with like resistors R1 and R2. The bridge is fed the horizontal diagonal; the vertical illuminates the entire width of the plate and a string of transducers is arranged on the opposite side of the plate, facing the stage T2, T4.

At the beginning, the bridge is balanced, the collectors of T2 and T4 are at the same potential, and the transistor circuit unbalance current I is zero. Now, let a very small DC current (i) pass through the moving coil; it will rotate it by a very small angle, enough, however, to reflect more light on P1 and less on P2, or vice versa. The conduction of the photoresistors varies accordingly, the transistor bridge is unbalanced, and current I, much greater than i, passes through the graphic recorder windings. There actually is current amplification.

Just as in the circuit described above, a calibrated resistor (R) is series-connected with the load (the recorder), which turns the amplifier into a closed-loop device. Current I, indeed, develops across R a voltage drop, v equals RI, which is applied to moving coil G in series with voltage, Vx, to be measured. Subjected to these two opposing voltages, the mobile coil rotates in a way to null the difference; i.e., Vx equals v. When this balance condition is reached, galvanometer current i is 0. In spite of the comparatively low values of G and R, the apparent input resistance is infinite (because of an inverse feedback-like effect). The system sensitivity can be modified, using for R a different value resistor.

AUTOMATIC DIMENSIONAL INSPECTION

Automatic inspection frequently applies to dimensional control or checking machine parts and other items. As the dimension to be monitored may rank anywhere from a microscopic whisker to a large beam, there is no unique method to be used; rather, many specific methods more or less convenient for solving a given problem are relied upon.

Take the case of steel plates leaving a steel mill. Suppose their width is to be inspected while the plates are carried by the conveyor system. It may be practical to test them in a vertical position with one edge pressed by the weight of the plate on the top of the supporting roller, forming the width reference line (Fig. 4-16). A spring-loaded feeler is pressed against the upper edge of the plate and its angular displacement is geared-up to move the slider of a potentiometer according to plate width. Thus, we have a variable-resistance (or resistance ratio) readout, serviceable for display or control purposes, or both.

This is an **analog** transducer. To provide for **digital** readout, it is necessary only to replace the potentiometer with

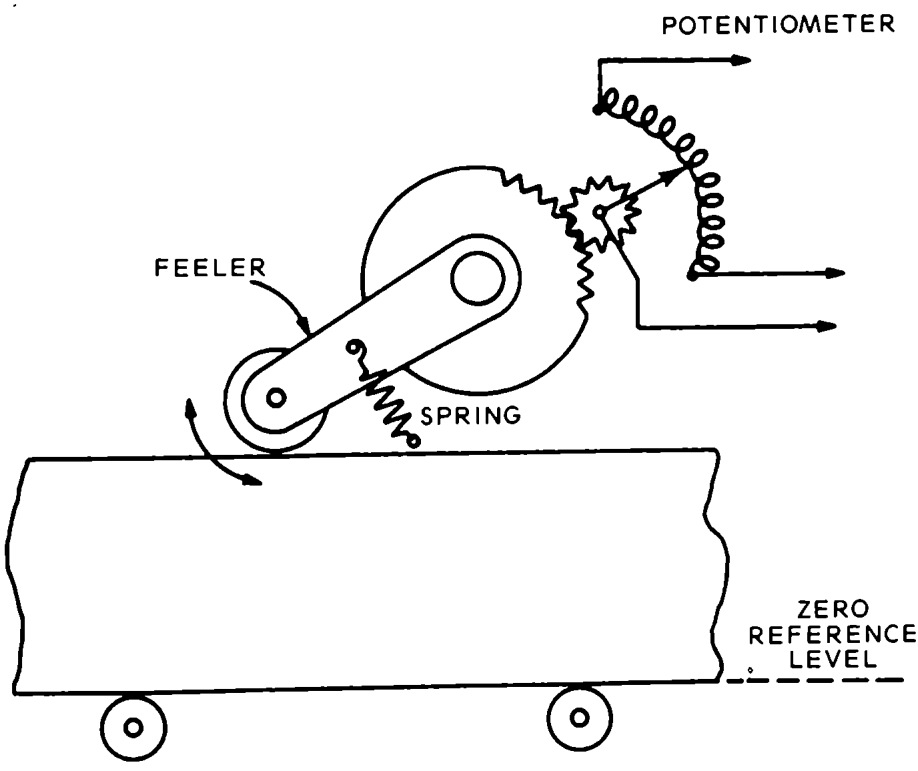


Fig. 4-16. Potentiometer steel-plate-width transducer.

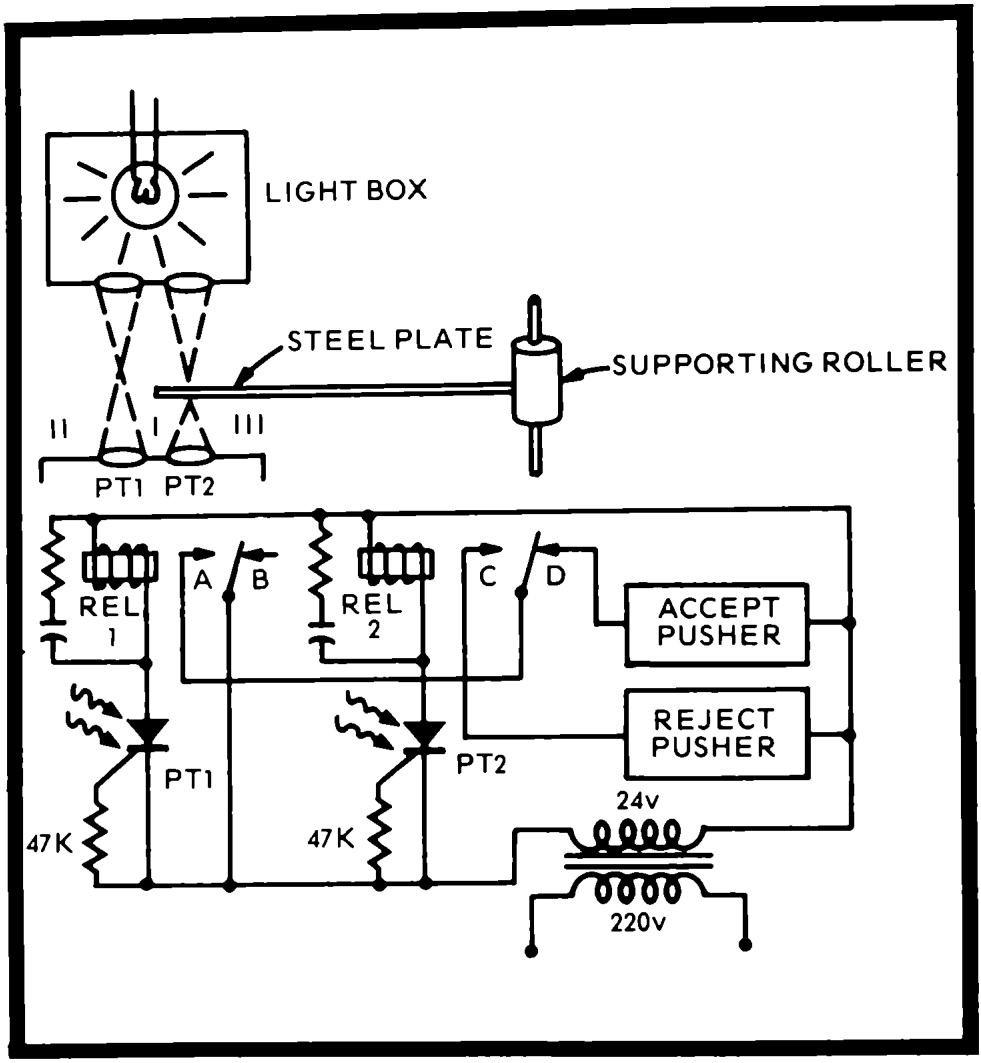


Fig. 4-17. Diagram of a photoelectric dimensional transducer.

an analog-digital converter (see Chapter 5) yielding a binary output. It might be even easier to measure the vertical displacement of the feeler arm by a (mechanical or electronic) counter.

Modern dimensional inspection devices preferably avoid physical contact with the item to be checked because of wear-down and jamming problems. The photoelectric device in Fig. 4-17 is satisfactory for the purpose. The lower plate edge is maintained at reference level as explained above and its upper edge runs between a light box and a photoelectric control box. The light emitted by a lamp is split in two parallel beams, concentrated by convex lenses to strike the light-sensitive junctions of two photothyristors, PT1 and PT2. For accurate sensing, the focal points of the beams are situated in the plate plane. TH1 and TH2 energize relays REL1 and REL2,

Steel Plate	Edge Position	Thyristor Fired		Contacts Closed	Sorting Decision
Too wide	II			b a	Reprocess
Too narrow	III	TH1	TH2	a c	Reject
Correct	I	TH1		a d	Accept

Table 4-1. Operation of the photoelectric dimension sensor.

respectively. The operation of this photoelectric control device is explained synoptically in Table 4-1.

According to the plate's width, its upper edge occupies positions I (correct), II (too wide, to be reprocessed for trimming) or III (too narrow, to be rejected), thereby intercepting one or both light beams. The light striking the window of a photothyristor fires it, energizing the accept or reject actuator to direct the plates to the warehouse or delivery batch, or to the scrap pile. The plates left by the pusher remain on the conveyor and are carried to the trimming shears for reprocessing. Once trimmed, they are inspected again.

AUTOMATIC QUALITY CONTROL

Here again, the number of possible defects is enormous, and the number of possible inspection methods is still greater. There always is a physical or chemical property presenting a characteristic change when the item is defective, and a transducer detecting this change is selected to issue accept-reject orders.

As an example, consider the task of inspecting steel plates for pinholes at the end of the milling line. This is important, for customers want plain plates, not sieves. (The rejected items can often be sold for definite applications at a reduced price.) As shown in Fig. 4-18, the rather simple control device uses photoelectric transducers. A transverse light source illuminates the entire width of the plate and a string of transducers is arranged on the opposite side of the plate, facing the light source. If there is a pinhole, it passes light to the corresponding light-sensitive device, preferably a photothyristor which energizes the valve of a paint pistol to mark the location of a pinhole. The best plates, obviously, are those showing no paint traces.

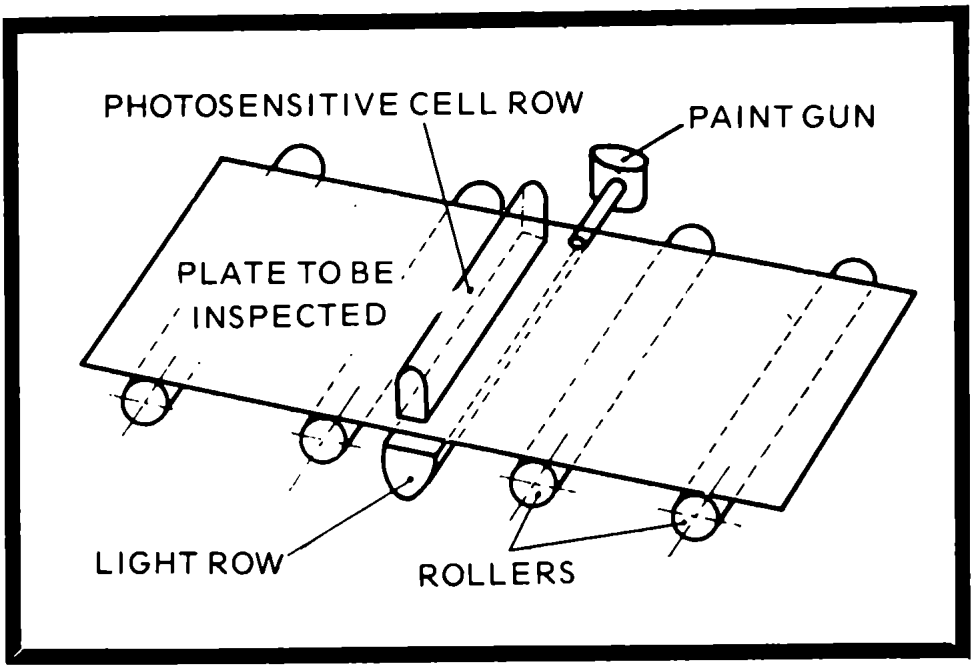


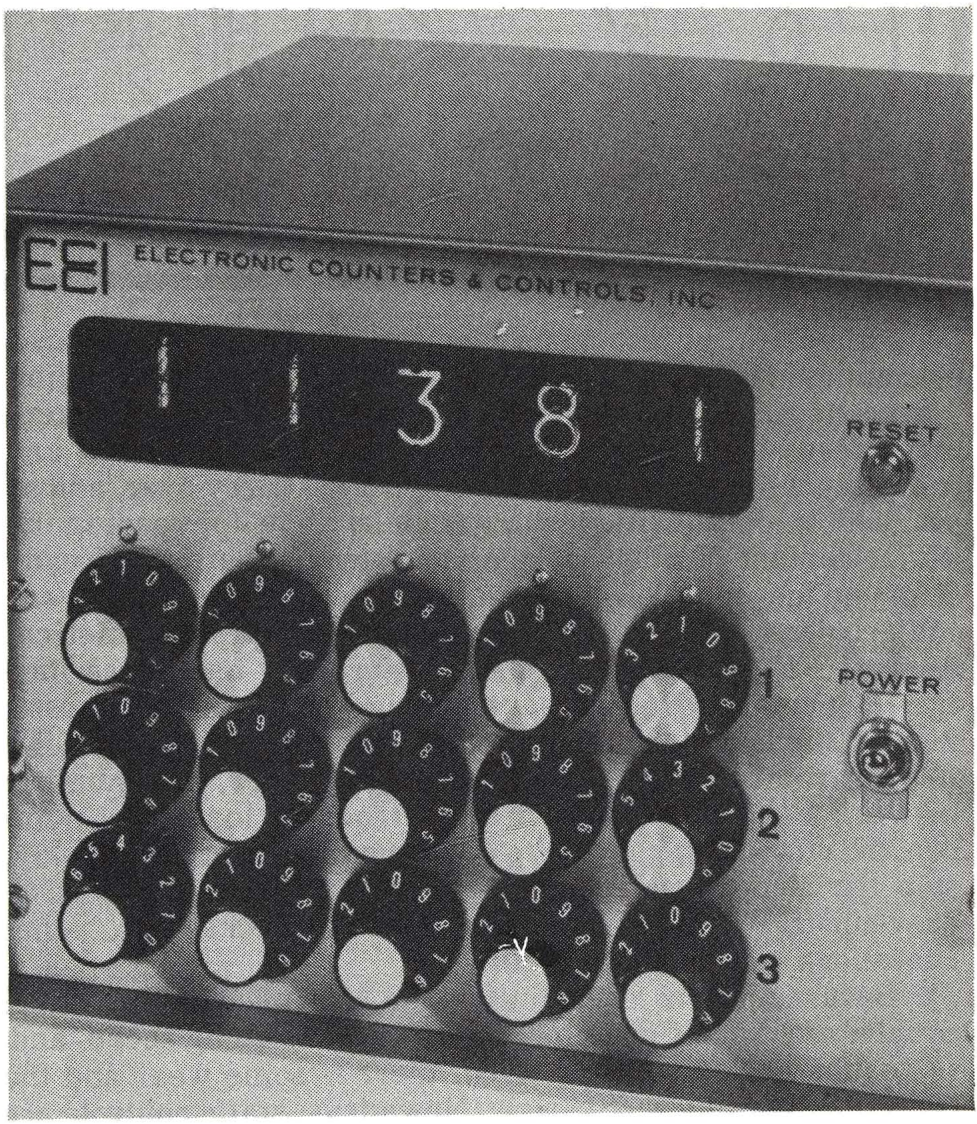
Fig. 4-18. Photoelectric detector used to spot pinholes in sheet metal.

In mass production operations, such as plate milling, information is often required on how frequently defects occur or on the percentage of rejects for statistical purposes, manufacturing process evaluation and cost accounting. Such information is provided by a counter controlled by a contact of the defect detector relay. This leads us to counting operations and devices.

ELECTROMECHANICAL COUNTERS

The "works" of an electromechanical counter are much like that of a stepping switch. When energized, an electromagnet attracts an armature, thereby pulling down a gathering pawl which is engaged with a 10-tooth wheel. The pawl rotates the wheel one tooth or one tenth of a revolution (Fig. 4-19). A set of rolls bearing the ten numerals from 0 to 9 are mounted coaxially with the toothed gear; they read, from right to left, units, tens, hundreds, and so on. While the units roll is mounted solidly with the toothed gear, the remaining rolls do rotate on the common spindle, though not freely. A small pin on the units roll drives the tens roll to the next figure when the units roll passes from 9 to 0; this is the mechanical way to operate the carry over. Again, the tens roll drives the hundreds roll to the next figure when passing from 9 to 0, and so on.

The counting magnet is operated by current pulses of suitable amplitude and duration, obtained preferably by a contact closure or from a thyristor. Many counters feature a zero-setting device operated by a distinct electromagnet, which allows for remote control. If such a zero-setting device is provided, **predetermined counting** can be obtained if a contact closes at the carry-over from 9999 to 0000 (in the case of a 4-figure counter). On the other hand, overcoming a slight amount of friction, the figure rolls can be rotated to introduce the complementary number. Thus, if 450 events are to be counted, the counter is preset to 10,000 - 450 or 9,550. Then, 450 events are needed to reach carry-over and close the contact.



Electronic counter designed for item counting, a coil winding control, batch counter, a cut-to-length controller, plus many other applications. Developed by Electronic Counters & Controls, Inc.

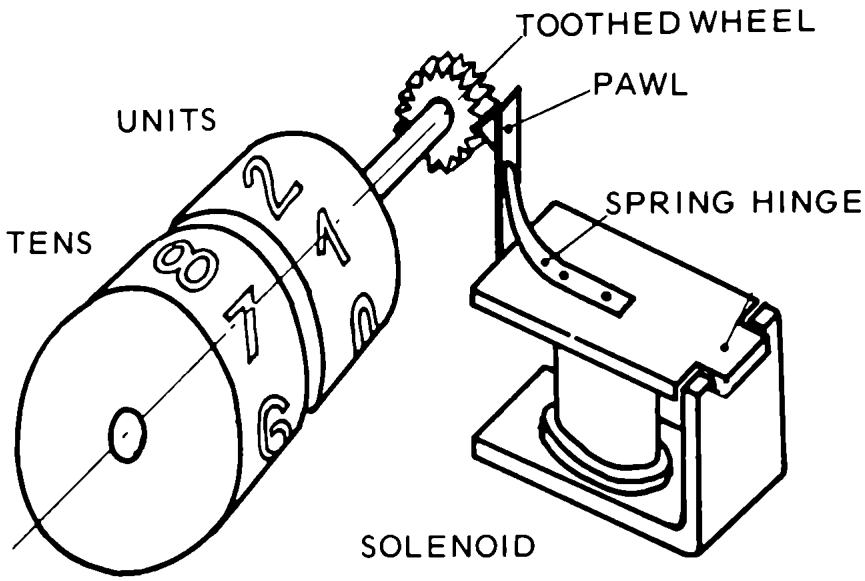


Fig. 4-19. Diagram of an electromechanical counter.

For special purposes, an electromechanical counter can be composed of as many 10-point steppers as there are digits. The count pulses are applied to the magnet of the units stepper. Processing from 9 to 0, this device closes a contact and energizes the magnet of the tens stepper for carry-over, and so on. Each stepper is fitted with a scale to allow for reading the actual contact position, so that the stepper battery already displays the count. But for more convenient reading, one contact deck can be used to switch some type of numerical display (Fig. 4-20). The homing feature has been described above; it can be wired so that the device comes to rest at 0000, with all the homing contacts series-connected. By means of a 10-point rotary switch, operated by a stepper, any number between 0 and 9999 can be preset. Reaching the preset number, the closed stepper and selector contacts energize a relay, stop the stepper rotation and signal and-or take action as convenient. If there still are contact decks left, other operations can be carried out, too.

Because of its operating mode, an electromechanical counter or stepping switch has a limited switching speed, and, since moving parts are involved, there is some wear and the possibility of failure, obviously. Nowadays, static controls are preferred, especially if utmost reliability is required. But the positive contact and easy switching diagram layout features of steppers may not be obtainable easily and economically by purely electronic means.

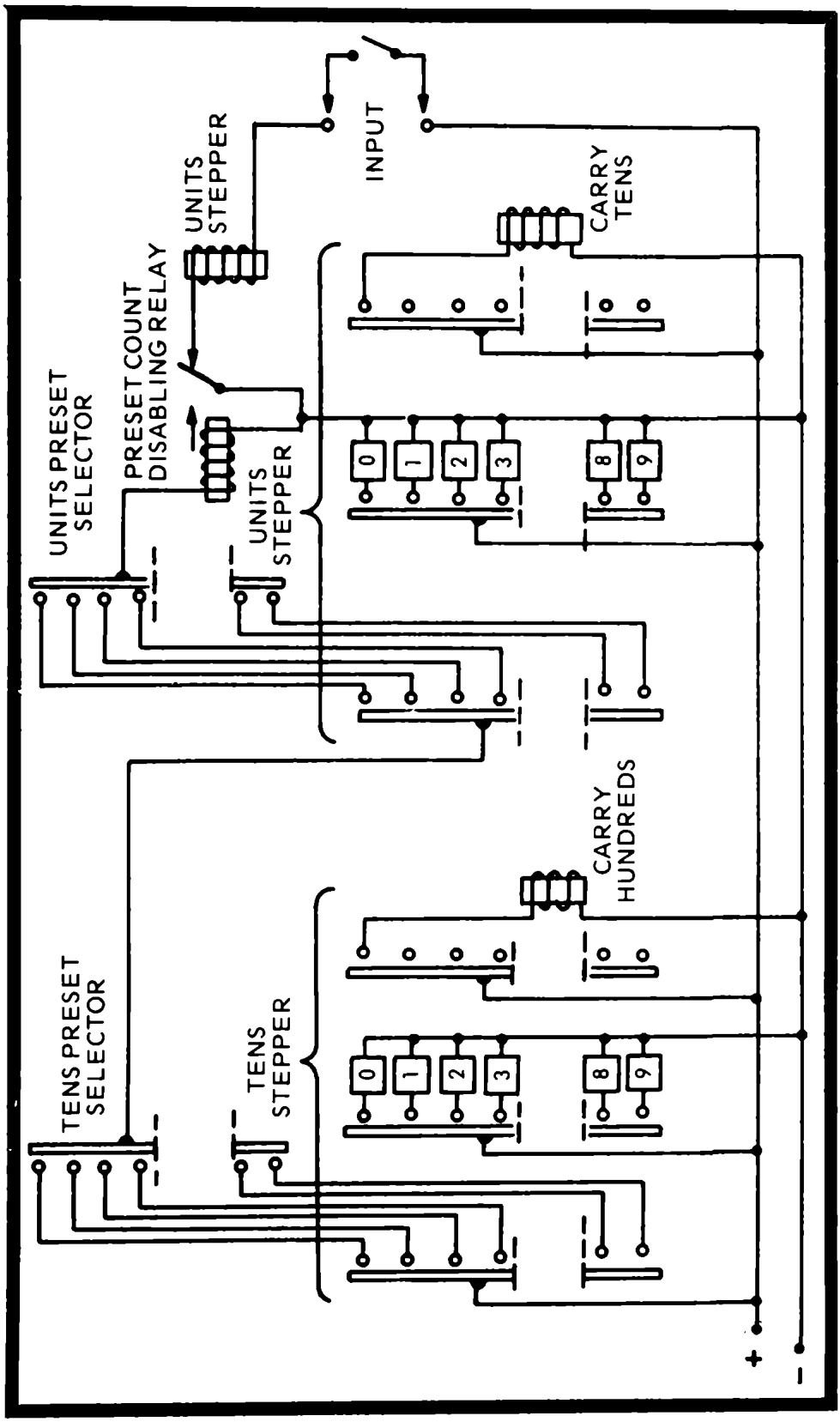


Fig. 4-20. Preset counter diagram using one stepper per digit.

ELECTRONIC COUNTING AND NUMERAL DISPLAY DEVICES

It is one thing to **count** the number of incoming pulses, i.e., develop a state in an electronic circuit corresponding to this sole number, and another thing is to **display** this number in a clearly readable way by decoding the electronic count. Before the event of semiconductors, counters used vacuum tubes, generally four twin triodes per decade; this explains the bulkiness of the Univac, the world's first electronic computer using thousands of tubes. So specially developed counting and display tubes appeared attractive, even if the count was not too easy to read.

The **Dekatron**, a cold-cathode gas tube, is such a combined counting and display device. A glass bulb contains a central anode made positive in respect to the 10 uniformly spaced, concentrically located cathodes (Fig. 4-21). Special transfer electrodes are arranged on both sides of each cathode; they are tied together in two Groups, A and B, with A on the left side and B on the right side of the cathode facing the anode. The counting pulse of negative polarity is applied directly to Group A, and through a delay line, which introduces a delay time of some microseconds, to group B.

When B+ is applied, one of the anode-cathode spaces ionizes, and the working cathode is surrounded by a glow visible through the glass wall; this identifies or "lights" the numeral of the corresponding cathode. The current through anode resistor Ra lowers the anode potential, preventing the glow discharge from reaching other cathodes. Because the transfer electrodes on both sides of the cathode are slightly positive in regard to the cathode, there is no reason for the glow to leave its cathode for another: the position, therefore, is stable.

A strong pulse applied to the input first reaches the Group A transfer electrode on the left side of the glow-surrounded cathode, and, since it presents a higher potential difference in regard to the anode than the cathode, the glow jumps over to it. Some microseconds later, the pulse appears on the Group B electrodes, and again attracted by the highest potential difference, the glow jumps to the next Group B electrode that, indeed is the right-hand transfer electrode of the next cathode in a clockwise direction. At the end of the pulse, all transfer electrodes return to a slightly positive potential; thus, it is the following cathode that attracts and retains the glow.

Hence, each pulse pushes the glow one step forward clockwise. Inserting the delay line in series with Group A instead of

B makes the glow run in the opposite direction, and there is a subtraction instead of an addition. By arranging a resistor (Rco) in series with cathode number 0, a carry-over pulse is obtained for advancing the tens-decade Dekatron by one position.

The **Trochotron** is another decade counter tube; however, it needs a separate display device. The Trochotron is a hot-cathode crossed-field vacuum tube akin to a magnetron. Its envelope contains, concentrically arranged around a central cathode, 10 groups of three electrodes: a **deflector** which shapes and positions an electron beam emitted by the cathode, a target that collects that beam, and a **grid** which switches it to the following position (Fig. 4-22). A cylindrical permanent magnet surrounds the tube to provide a magnetic field that works with the static electrode field to switch the beam device.

With the application of B+, all deflectors are at the same positive potential, but the magnetic field prevents the electrons from reaching the electrode groups, and there is no beam. Reducing the potential on one deflector by applying a negative-polarity pulse results in the formation of an electron

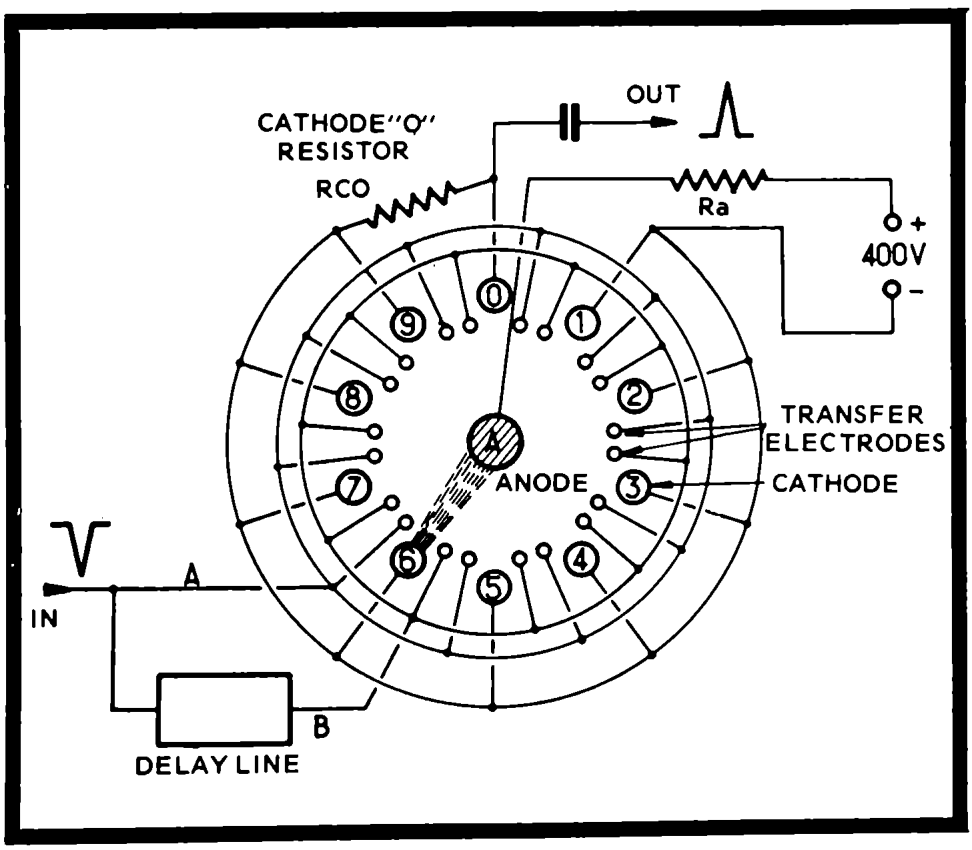


Fig. 4-21. Drawing of the electrode arrangement in a Dekatron tube. The position of the glow identifies the digit.

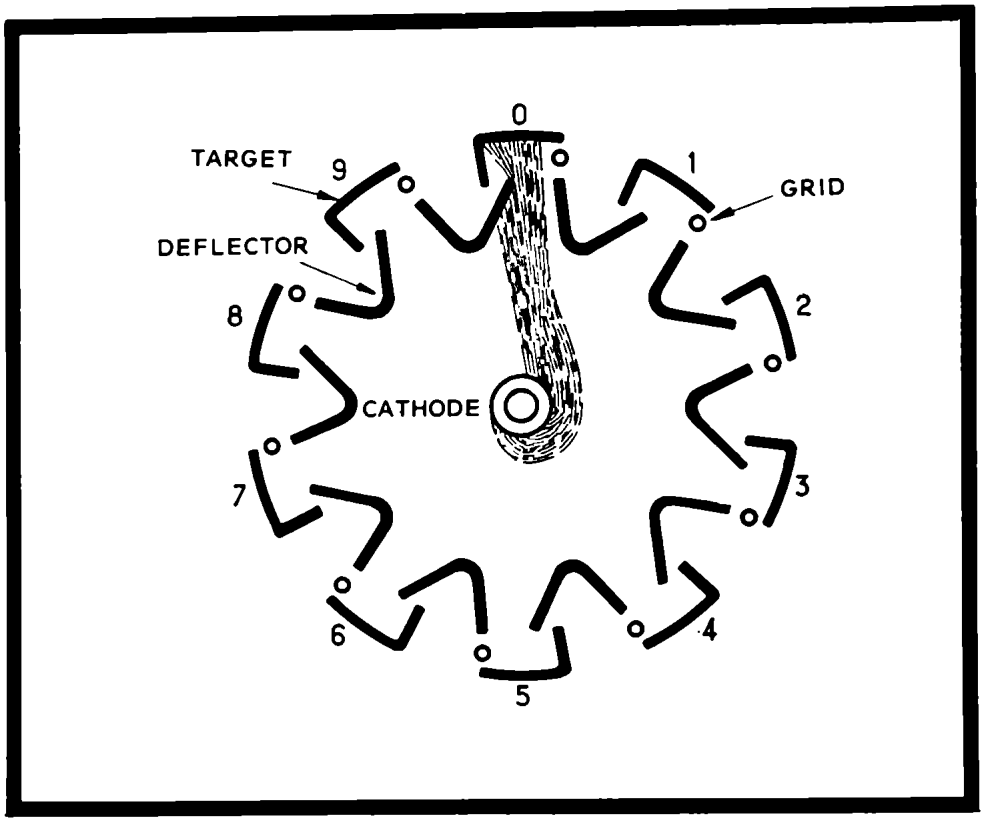


Fig. 4-22. Diagram of the element placement in a Trochotron tube.

beam toward the electrode group associated with the deflector concerned as shown on the diagram, and the corresponding target passes a current. The beam is retained in this position. A pulse lowering the grid potential by about 15v will switch the beam to the next position. Since only adjacent grids are able to act on the beam, odd- and even-order grids are tied together internally to form two groups, A and B, fed with opposing waveforms produced by an Eccles-Jordan circuit. Crossing connections A and B will cause the tube to count backward instead of forward. A resistor inserted between each target and B+ provides an output on the current-carrying target. Using a Nixie-type display tube (next paragraph), it is necessary only to connect each numeral-shaped Nixie cathode to the corresponding Trochotron target to display the numeral whose target is carrying current, suitable Nixie electrode voltages provided, of course.

IN-LINE NUMERAL DISPLAYS

Though an ingenious concept, the Dekatron offers a display type that does not make for immediate effortless

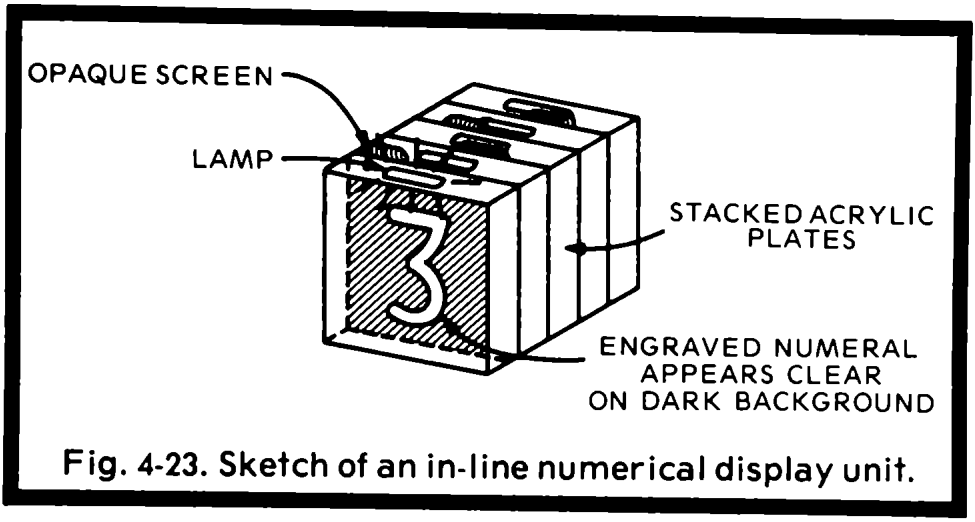


Fig. 4-23. Sketch of an in-line numerical display unit.

reading (especially when large numbers are involved) because the glow spot position is to be interpreted as a numeral, and then the numerals composing a number do not appear on a common baseline. For the sake of quick and reliable reading direct-reading in-line displays are now superseding all other readouts. In some applications, however, such as working with radiations where counters are operating all day long, there is no point in reading actual numbers; it is rather the rythmn of flickering lights that provides a feel of the increase or decrease in radiation intensity.

The now familiar Nixie-type glow display tube is an assembly of 10 numeral-shaped wire cathodes numbered from 0 to 9. The cathodes are superimposed and mutually insulated, and arranged before a common anode. On the application of about 100 volts between the anode (tied to +) and one of the cathodes, the corresponding wire numeral stands out glow-surrounded, whether it is in the first or in the tenth plane. Since the cathodes are superimposed, all numerals appear at the same place. The Nixie tubes composing a readout are lined up to provide for convenient in-line reading. Recently, several manufacturers have undertaken the production of complete glow-type display units containing a number of Nixie electrode packs arranged side by side in a common reactangular glass envelope, thus reducing the spacing between numerals and simplifying assembly operations.

An optical counterpart of this electronic display is shown in Fig. 4-23. The 10 numerals are engraved in transparent acrylic plates individually edge-lighted by tiny lamps. These plates behave as light conductors: the light enters by the edge and can leave only at the plate via the sloping sides of the engraved track. If a lamp is lit, the corresponding numeral is illuminated on a dark background.

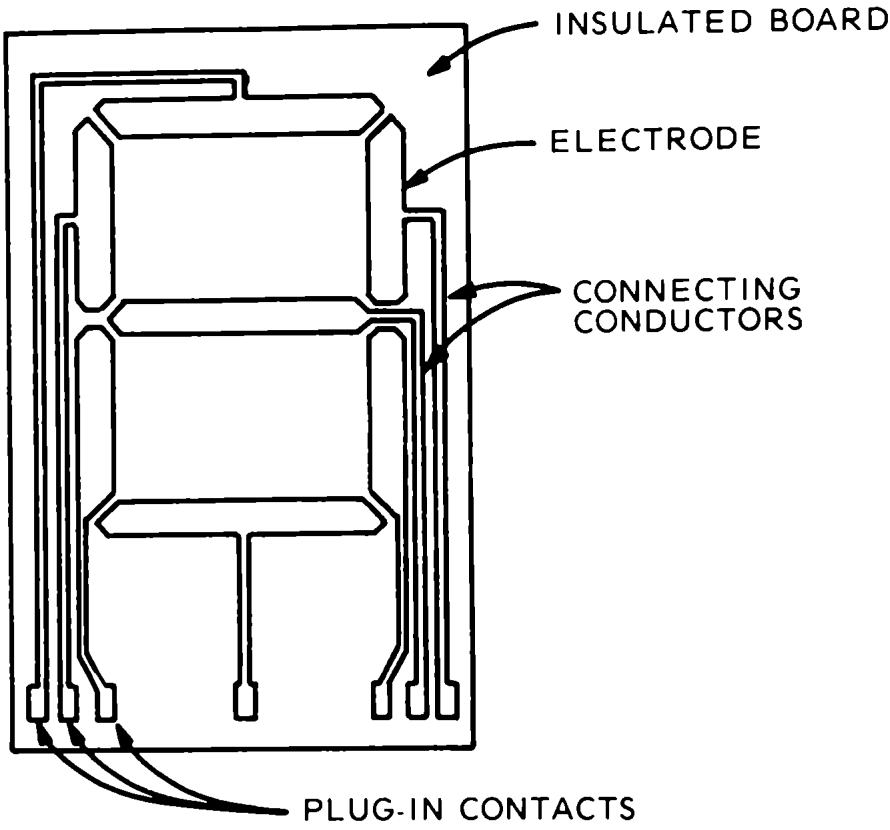


Fig. 4-24. Drawing of a 7-segment numeral display.

Another type of numeral display is composed of seven segments located in a single plane (Fig. 4-24), each containing the basic elements for forming each of the 10 numerals by selecting the suitable segments by means of an encoder. Fig. 4-25 shows the form of the numerals thus obtained. While their aesthetics are debatable, the ease of reading is obvious, and one rapidly gets familiar with it. (The magnetic ink identification of checkbooks does not feature pleasant numeral shapes either.)

Such a display type can be operated by any method that allows for selective illumination of each of the seven segments. This effect can be obtained optically by a lamp arrangement behind an opaline glass with opaque separations delimiting partitions corresponding to the segments to be illuminated selectively. In this case, a small lamp is used to illuminate each of the segments. While this device is easier to build than the acrylic pack in Fig. 4-23, it needs a decoder.

However, recent electronic technology soon will allow for the economic mass production of such 7-segment displays

using electron-luminescent materials. There are phosphors that become luminescent upon the application of an electric field. On the insulating base plate in Fig. 4-24, the indicated segment and connections pattern is deposited or it is etched-out of a copper-plated insulated base by printed-circuit techniques. Then, the electroluminescent material is deposited on the segments. Each segment finally becomes a transparent gold layer used as the second electrode of a capacitor. Such a segment emits a well visible greenish light when fed an exciting voltage.

Another, still more recent technique uses gallium arsenide-phosphide injection diodes which are the shape of the segments. These diodes emit a red light on application of 4 to 7 volts DC. Hitachi is testing such a display powered by a 1.75v source; its current drain is 50 ma. The brightness is said to be 120 foot-lamberts, with a buildup time of 1 ns.

Maybe liquid crystal displays will be with us some day. This refers to a peculiar liquid, normally transparent, that turns opaline-milky when subjected to an electric field. On a dark background, the liquid in its milky state appears the brighter the higher the ambient illumination level. Up to now, such devices have not been marketed.

PRESET ELECTRONIC COUNTERS AND DIGITAL TIME-DELAY GENERATORS

The carry pulse of the highest power decade of a counter is serviceable for signaling purposes and for triggering an ac-

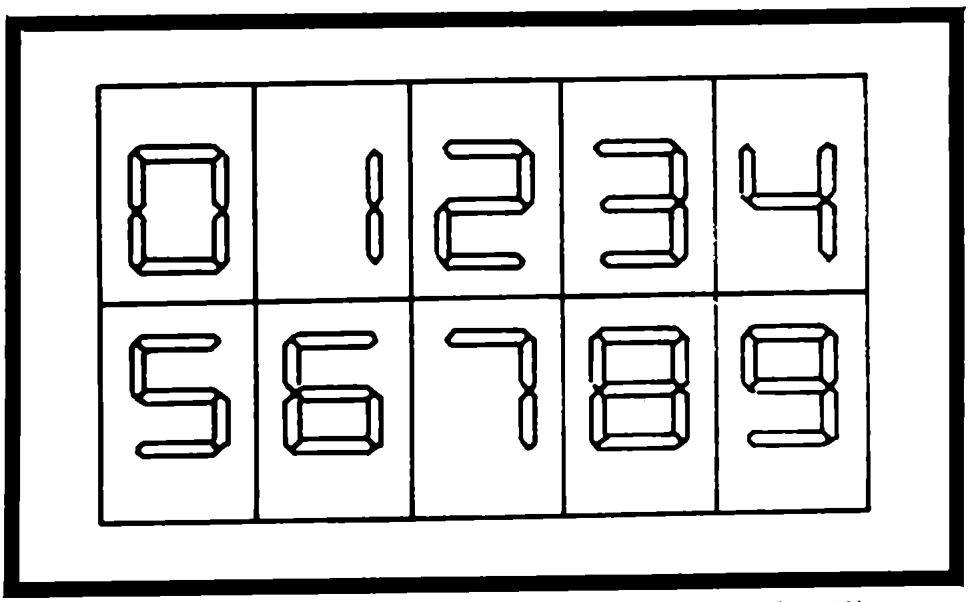


Fig. 4-25. Drawing of the numerals obtained with a 7-segment device.

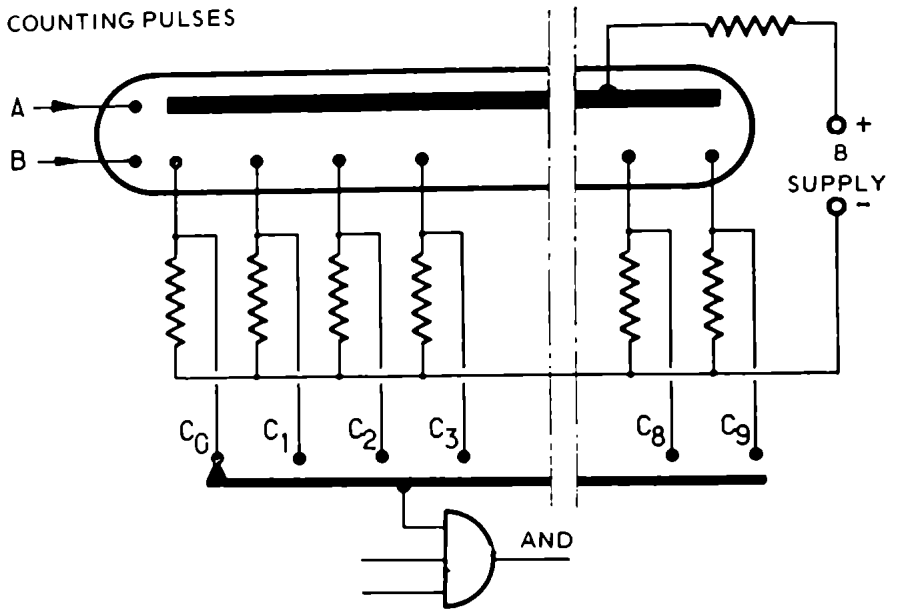


Fig. 4-26. Schematic of a Dekatron count preset stage.

tuator, after filling, say, 100 pills in a box, using a 2-digit counter. Most electronic counters, however, can be preset to deliver the control pulse required after any number of input events. In Dekatron and Trochotron circuits, load resistors are series-connected with the 10 targets or cathodes, and, of course, a signal develops across a resistor tied to a current-carrying electrode.

To derive a preset counter using Dekatron decade stages, for instance, cathodes CO to C_9 are tied to B- through series resistors, and a 10-point selector switch can be connected to the cathodes. Consequently, a signal will appear at the switch wiper if that cathode is operating. Each digit requires a Dekatron and a selector switch, and the wiper outputs of these switches are combined by an AND circuit yielding an output only if all of the possible input signals are present (Fig. 4-26). This output is the control signal corresponding to the preset count.

Using a preset counter, any time delay can be generated with a high degree of accuracy by counting the pulses derived from a high-stability crystal oscillator. Take a pulse repetition frequency of 100 kHz; two succeeding pulses, then, are separated by 0.01 ms. To generate an accurately timed time interval of 328.54 ms, it is necessary only to count 32,854 pulses.

Looking at the functional diagram of such a digital time interval generator in Fig. 4-27 makes the term "it is necessary

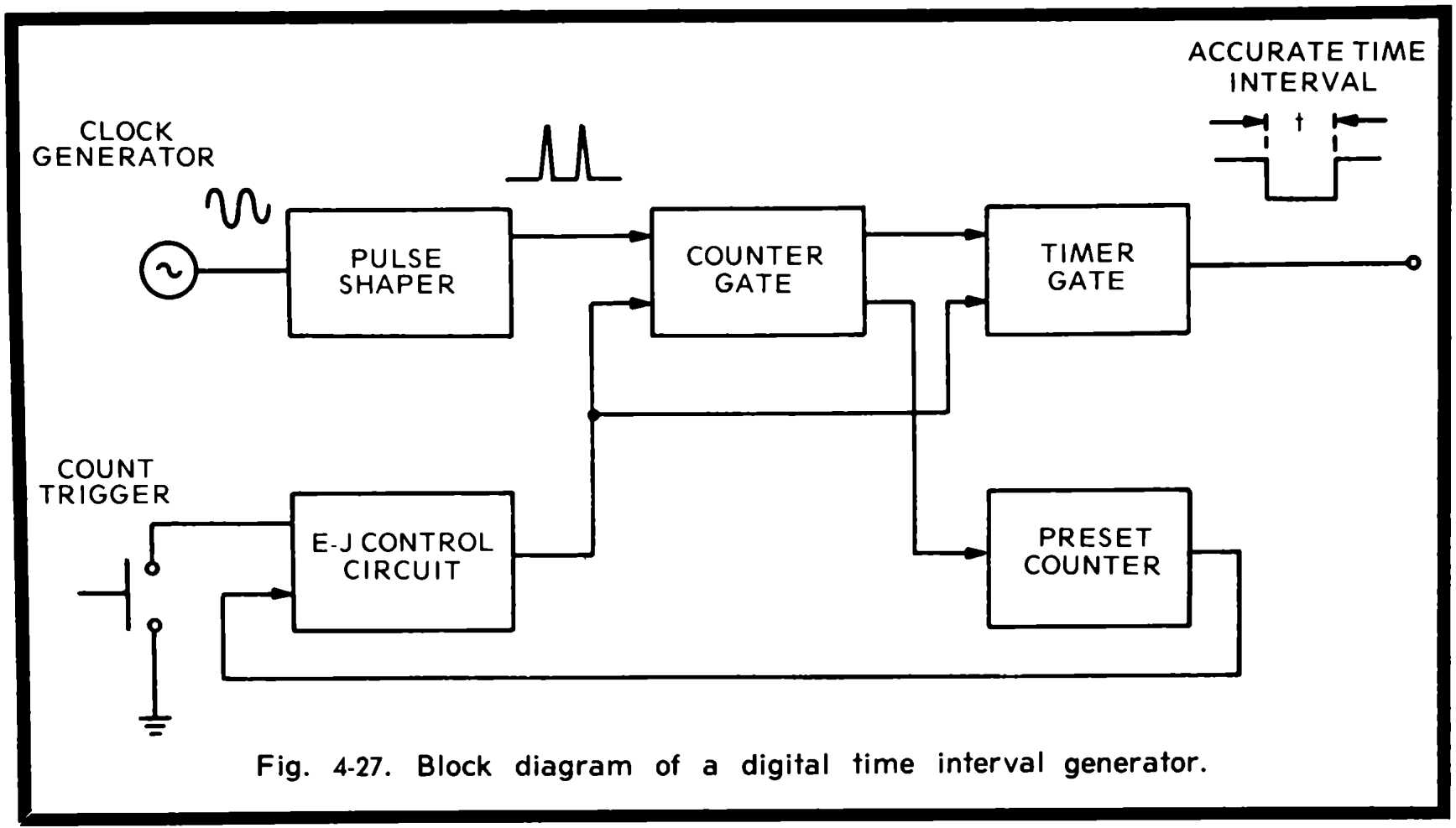


Fig. 4-27. Block diagram of a digital time interval generator.

only" an obvious understatement. A crystal-controlled "clock" generator delivers a high-stability CW signal which a waveshaper transforms into clock pulses of same stability. The pulses are applied to the counter and timer gates normally inhibited by the scale-of-two control circuit. Closing a contact as indicated or injecting a triggering pulse tips over this circuit and thereby opens the two gates. While the timer gate initiates the time delay to be generated, the counter gate starts the preset counter. On reaching the preset count, the latter issues a control signal, flipping the scale-of-two back to its initial state. Closing the timer gate ends the time interval.

With a 100-kHz clock frequency, two succeeding pulses are separated by a 10-microsecond time interval; this also is the resolution of the counter, limiting the definition of the generated time interval. Using a 1-MHz clock allows for a 1-microsecond resolution. This means that any whole number of microseconds can be generated; the actual accuracy of the preset interval, however, depends not only on the pulse count precision, but also on the accuracy and constancy of the clock frequency.

This complex while highly accurate timing device is a typical digital solution of a problem that also can be solved by using an analog method such as the charge or the discharge of a capacitor. This leads to a comparatively simple circuit, even if refinements such as constant-current charging are used. But the resolution of such an analog timing device remains limited.

EVENTS-PER-UNIT-TIME COUNTER

Fitting a counter with a gate passing pulses during an accurately defined time t (such as one second) allows for counting the number of events per unit time (EPUT). The interval generator circuit in Fig. 4-27 can be used to open the gate during time t. A 1-second time interval is obtained which counts 100,000 pulses spaced at 10 microseconds (clock, frequency 100 kHz). Losing one pulse would result in only a 0.001 percent error.

Such an EPUT counter can evalutate every phenomenon, convertible to voltage or current pulses, developing during a given time t. To measure the unknown frequency of a CW wave, the signal is first applied to a pulse former (a Schmitt trigger), and the resulting pulses are applied to a normally inhibited counter gate. On initiating the time interval, the gate opens and passes the pulses to a decimal counter (Fig. 4-28). After time t (preferably one second), the gate closes, stopping the counter. Provided t equals 1 s, the count displayed equals the unknown frequency in Hz or cps.

To accurately measure a time interval (for camera shutter performance evaluation, for instance), the time interval generator is replaced by a pulse former that controls the counter gate according to the signal of unknown duration, and a clock frequency generator is connected to the fx terminals.

An EPUT counter is also serviceable for evaluating the rotating speed of a machine using a suitable transducer (Figs. 2-32 or 2-33) producing a pulse per revolution. Since this magnitude is conveniently evaluated in turns per minute (tpm), a direct reading is obtained for a 1-minute counting interval. If this measuring time is deemed exaggerated, a transducer delivering a number of pulses per revolution will reduce it proportionally.

The flow of a fluid within a pipe is often monitored by a small turbine located in the pipe and rotated by the fluid stream. A small magnet, fitted on the turbine periphery, induces a flux change in a coil arranged on the outside of the pipe (to be made of a nonmagnetic material), so that a voltage pulse is generated per turbine revolution. These pulses are counted as before for a direct fluid flow reading. Many similar applications can easily be imagined.

DIGITAL VOLTMETERS

The conventional d'Arsonval-type voltmeter with a pointer moving in front of a calibrated scale is an analog device requiring some mental interpolation for reading. Furthermore, its reading accuracy is necessarily limited to about 1 percent or a fraction thereof. In many applications,

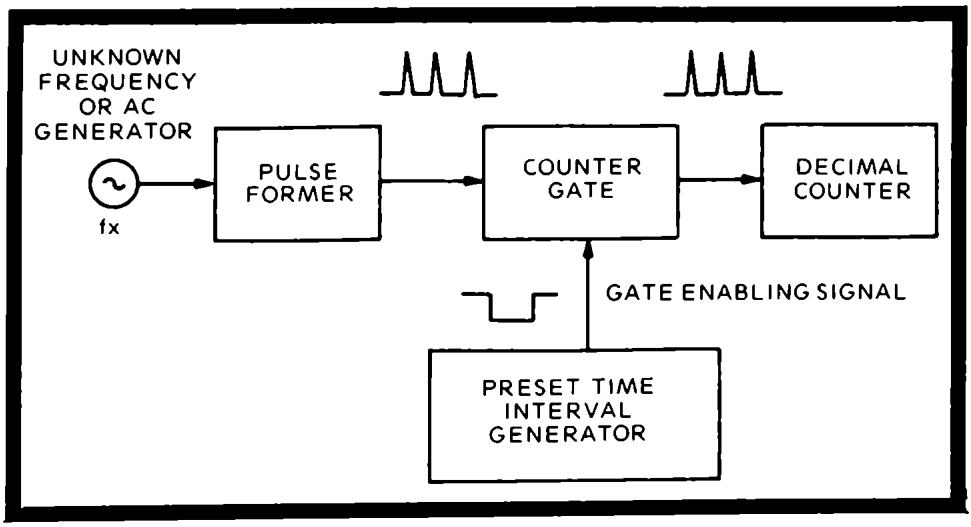


Fig. 4-28. Block diagram of an events-per-unit-time counter.

these instruments are now increasingly superseded by **digital voltmeters** which effortlessly read 3-, 4- or 5-digit numbers displayed by Nixie tubes or similar numeral display devices. Beside the quick and practically error-free reading feature, formerly unheard-of measuring accuracies are obtained. Because of their complexity, DVMs evidently are much costlier than d'Arsonval meters; but for laboratory or production line work, they may actually be profit-earning due to speedier readings and reduced errors. And using integrated circuits, DVMs are no more bulky, necessarily, than conventional VOMs.

The operating principle of a DVM is illustrated by the diagram in Fig. 4-29. Using the potentiometer method, the unknown voltage, Vx, is compared with a variable reference voltage, a perfectly linear sawtooth voltage produced by a ramp (voltage) generator (waveform A in Fig. 4-30, showing the various operating waveforms above a common time scale). Notice that ramp voltage A varies from +12v to -12v during 60 ms in this particular Hewlett-Packard DVM.

By means of a manually operated or automatic decade amplifier and-or attenuator, Vx is given an amplitude between +12v and -12v, suitable for comparison with the reference voltage, and both voltages are injected into a differential amplifier which compares them and issues a coincidence pulse m (wave B) at the very instant Vx crosses the ramp (Point M). In the case of an unknown voltage, -Vx, of negative polarity, a coincidence pulse n would occur at Point N. On the other hand, a second differential amplifier supplied with the ramp voltage issues a zero-crossing pulse p (wave C) at the very instant (P) the ramp crosses the zero amplitude level. Thus, a voltage-time conversion is obtained, for time interval t, between M and P (or P and N), is strictly proportional to Vx of whatevery polarity.

Pulses m and p (or p and n) open and close the counter gate (wave D) so that during a time interval t, the clock pulses delivered by a crystal-controlled clock frequency generator are passed along to the decade counter (wave E), displaying the count by means of Nixie tubes or similar devices. With a pulse repetition frequency of 400 Hz and a ramp slope of 400v per second, each counter stroke is worth 1mv at the voltage comparator input. The settings of the decimal amplifier-attenuator define the actual position of the decimal point. A polarity detector determines the sequence of arrival of the coincidence pulses (i.e., m before p, or p before n) and energizes the sign + or - accordingly.

While the pointer of an analog meter follows the fluctuations of the unknown magnitude, a DVM in a way "shoots a

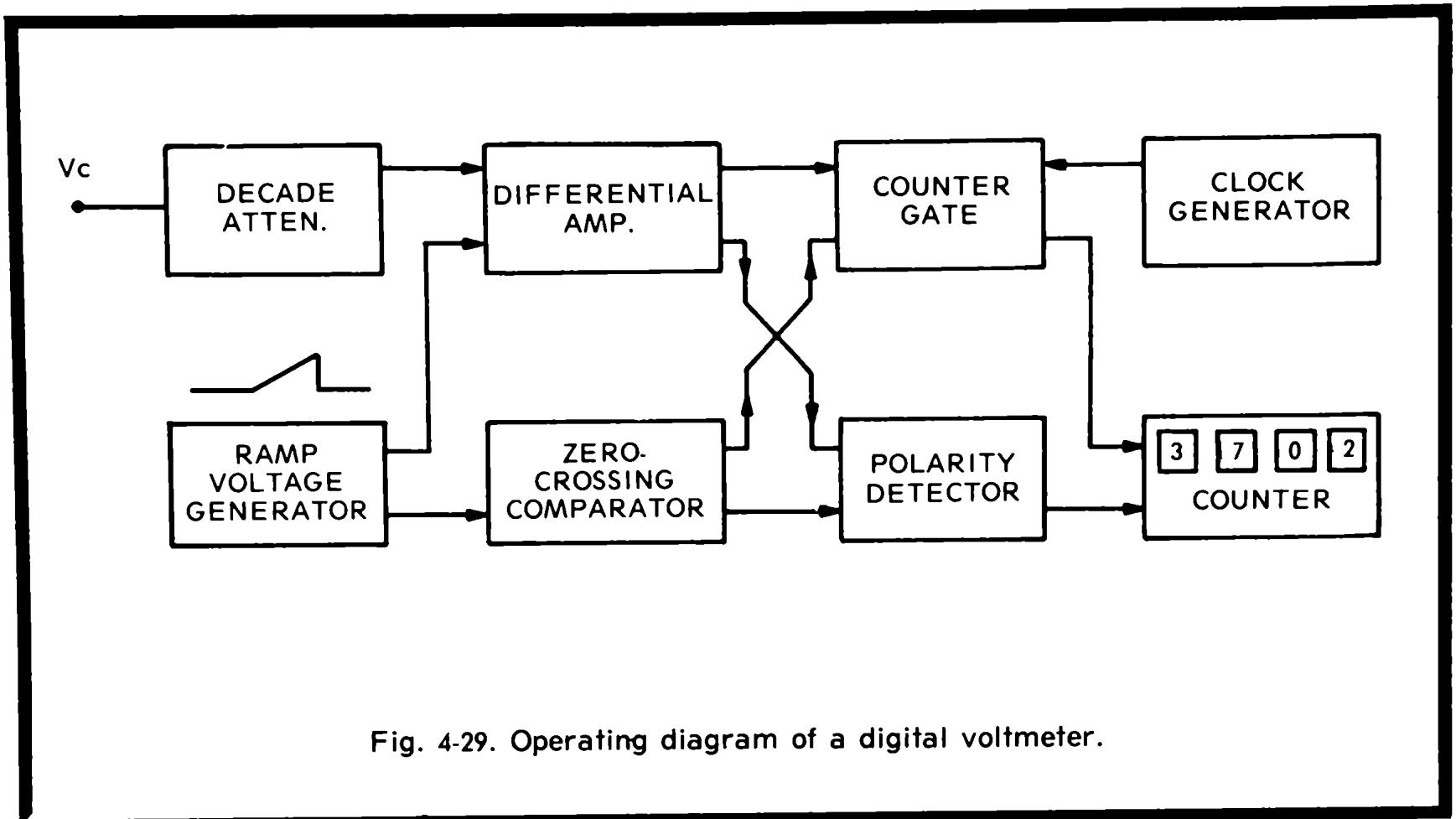


Fig. 4-29. Operating diagram of a digital voltmeter.

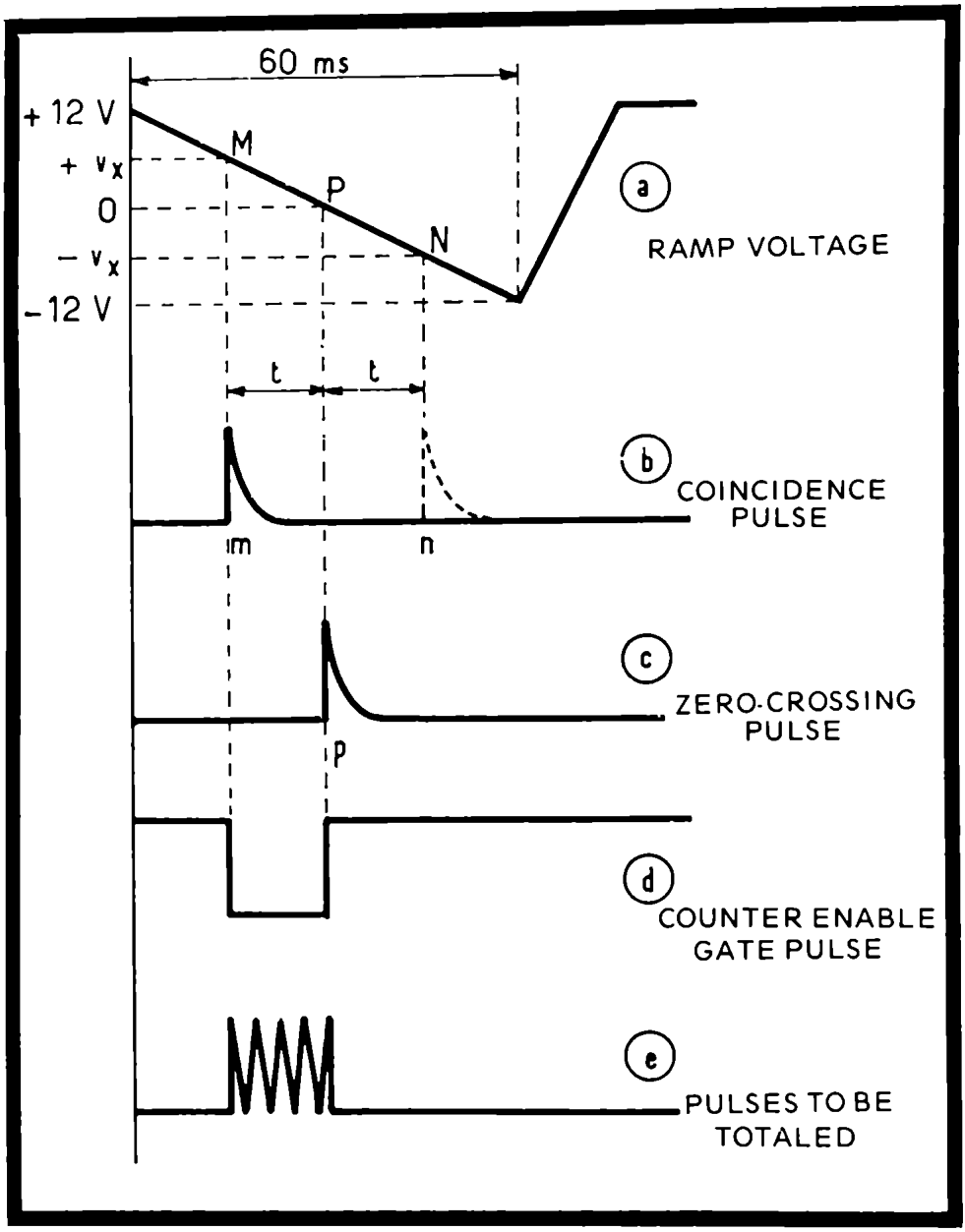


Fig. 4-30. Waveforms encountered in the DVM diagrammed in Fig. 4-29.

picture" of the situation at a given instant. A sampling device initiates new measurement processes repetitively at intervals variable between 0.2 and 5 s (it takes 0.2 s to complete a measurement). The model described allows for measuring DC voltages from 0.001 to 99.9v with an accuracy of plus or minus 0.05 percent plus or minus 1 digit. With little additional circuitry, currents can be measured, too, as well as AC volts (peak or RMS), ohms and other electrical magnitudes. Signals for operating a print-out device are easily provided.

A DVM has many uses in industrial electronics, too. Previously, we discussed the measurement of nonelectric magnitudes by pulse counting. Different techniques may be used, too, such as accurate DC voltage measurements. For instance, the speed of a rotating machine can be monitored by coupling with a tachometer generator whose speed-proportional output voltage is compared with the ramp voltage. And most resistive transducers (thermistors, photoconductors, variable resistors) as well as voltage- or current-generating transducers (thermocouples, photovoltaic cells) yield a voltage related to the magnitude under control that can be compared to the ramp voltage. The number displayed may not directly read that magnitude, a temperature for instance. This does not matter if the operator knows the reading to be looked for and the limits not to be exceeded.

BINARY COUNTING

The decimal notation we are accustomed to using is based on the ten digits from 0 to 9, comparable to our ten figures (digit or finger). Though the 10 numerals are easy to write and to print, there is no electric or electronic device available

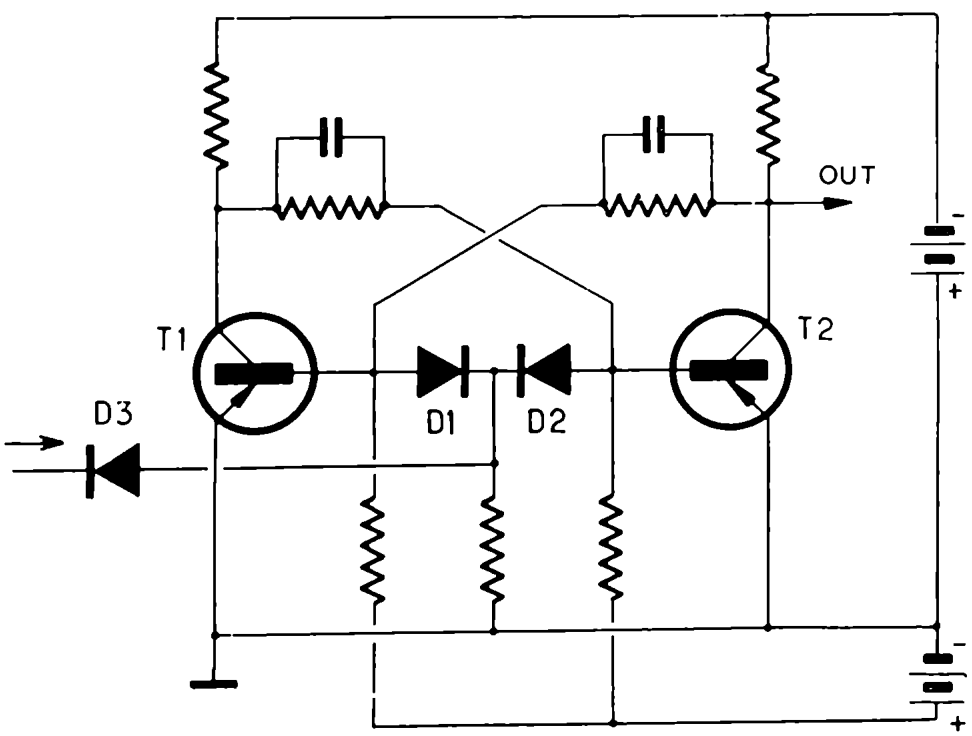


Fig. 4-31. Schematic of an Eccles-Jordan binary counter stage.

n	2^n	terms retained	decision	in binary
10	$2^{10} = 1{,}024$	1,024	yes	1
9	$2^9 = 512$	512	yes	1
8	$2^8 = 256$	256	yes	1
7	$2^7 = 128$	-	no	0
6	$2^6 = 64$	-	no	0
5	$2^5 = 32$	32	yes	1
4	$2^4 = 16$	16	yes	1
3	$2^3 = 8$	8	yes	1
2	$2^2 = 4$	-	no	0
1	$2^1 = 2$	-	no	0
0	$2^0 = 1$	1	yes	1
		1,849		11100111001

Table 4-2. Digital terms composing the decimal number 1,849.

featuring the corresponding 10 stable states with good reliability at an acceptable price. (The Dekatron and Trochotron tubes descirbed above meet these conditions, but are less convenient to use than modern in-line numerical readouts.) The binary numeration system contains but two numerals, 0 and 1, which can be assigned to the two operating states of a good many technical devices such as open-closed valves, on-off switches and relay contacts, and start-stop motor controls. When Hamlet exclaimed, "To be or not to be, that's the question!" he stated that indeed there was a choice between the two possible options of a truly binary alternative.

In the electronic circuit domain we are primarily concerned with, all active devices (tubes, transistors and thyristors) can be switched between two well-defined states, a virtually nonconducting and a heavily conducting one. Thyratrons and thyristors cannot even be operated by graduate variation, for only conducting and de-energized states are possible. Switching transistor operation falls into the same category. Opposing these switch-operated components are the 2-state devices also examined in Chapter 3. The **multivibrator**, a free-running relaxation oscillator, obviously has no stable position; the **flip-flop**, used to generate time delays, offers one stable state to which it returns some time after having been tipped over; the **Schmitt trigger** features two input level-dependent states; and the **Eccles-Jordan** or **scale-of-two** circuit is a pulse-switched device with two stable states. Since this is the basic component of an electronic counter, it is discussed to some detail later. But first, the fundamentals of binary computation must be understood.

The decimal numeration we are accustomed to is based on the number 10 and the binary system is based on the figure 2. This means that any number can be expressed by a binary number that is a sum of powers of 2. For instance, take terms composing number 1,849 in Table 4-2. In this decreasing sequence of powers of 2, we can write only the decision (YES or NOT) to retain or not to retain each term of the sum; in binary language, 1 stands for YES, and 0 for NO. Thus, the decimal number 1,849 becomes 11100111001 in binary notation. Of course, the digits from right to left designate an increasing sequence of exponents starting with 0. The greater the number, the higher, obviously, is the last (left-hand) exponent encountered, the so-called "weight" of the digit. You will notice the decimal 1,849 is a 4-digit number, while its binary counterpart is made up of 11 digits; but to get any 11-digit binary number, you need only 22 numerals, while it takes 40 fonts to make up any 4-digit decimal number.

Binary arithmetic has very simple rules conceived by the German philosopher Leibnitz about two centuries ago:

0 + 0	0	0 x 0	0
1 + 0	1	0 x 1	0
1 + -	10 (!)	1 x 1	1

So we can correlate decimal and binary numbers as shown below:

Decimal	Binary	Decimal	Binary
0	0	7	111
1	1	8	1000
2	10	9	1001
3	11	10	1010
4	100	11	1011
5	101	12	1100
6	110		

It is easy to check:

5	101
+7	111
12	1100

On the other hand, take 3 x 4:

11 x 100

00

00

11

1100 corresponding to 12.

Getting back now to the mechanism of binary electronic counting, notice that at the beginning, all the counter stages are nonconducting (0). The first incoming pulse tips the last stage over to the conducting state (1); the next pulse flops it back to nonconduction (0), while transmitting a carry pulse to the second stage, throwing it into conduction (1), so that the counter now reads 10, and so on. It is that easy.

ECCLES-JORDAN COUNTER DECADE

The fundamental scale-of-two circuit in Fig. 3-51 is often fitted with diodes to securely switch the counting pulses to the

receptive transistor base (Fig. 4-31). This circuit uses negative polarity control pulses (or, alternately, positive pulses with NPN transistors). Diodes D1 and D2 are biased to transmit the switching pulse only to the base of the blocked transistor, and diode D3 blocks any possible positive pulses capable of deceiving the counter.

Consider T1 conducting and T2 nonconducting to begin with; call this stage 0. The first pulse passed by D3 can act only on T2 via D2 (D1 is biased to nonconduction), T2 is thrown into conduction, the scale-of-two tips over, and T1 is blocked. This is state 1. There are no further possibilities for this bistable taken alone, for the next pulse will return it to the original 0 state; but the pulse input of a second scale-of-two can be connected to the output of the first so that a carry pulse is injected into the second stage each time the first one flips from 1 to 0. Hence, four pulses are needed at the input of the first E-J to get one pulse at the second E-J output.

Schematically representing each E-J by a rectangle divided in two halves (standing for the switching elements), Fig. 4-32 shows four cascaded E-Js composing a "scale of 16," for 16 input pulses are needed to get one output pulse. The device actually carries out a **frequency division**: each stage dividing by 2, a division by 2^n is obtained with **n** stages. Here, **n** equals 4, so that 2^4 equals 16. This principle allows for dividing the frequency of a clock oscillator while maintaining rigid phase relations. (Frequency division can be attained, too, by synchronizing a free-running multivibrator with the clock oscillator, but then some phase shift cannot be avoided.)

The switching pulses need not necessarily be of the short-duration type; more or less perfectly rectangular waves can be used as well, as shown by Fig. 4-33, illustrating the waveforms encountered in a divide-by-eight circuit using three cascaded E-Js. It is always the wave front that triggers the tip over, and a steep rise is required for reliable operation; the wave tail does not matter. For best results, the trigger

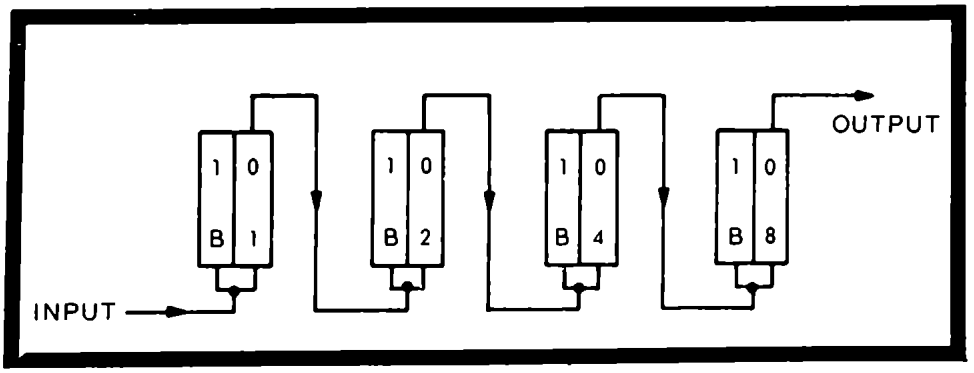


Fig. 4-32. Block diagram of a 4-stage E-J making up a "scale of 16."

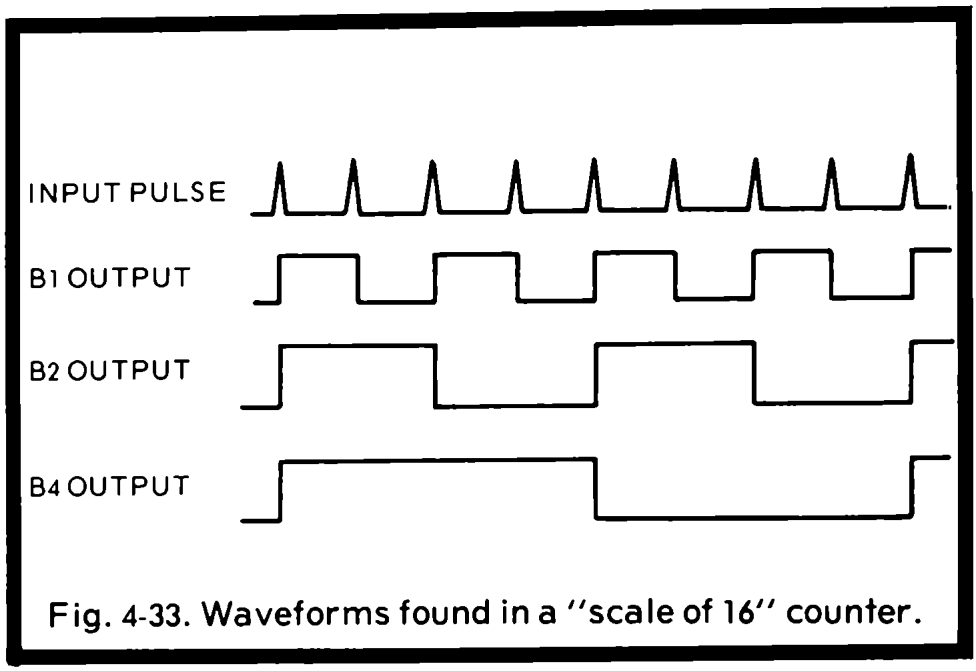


Fig. 4-33. Waveforms found in a "scale of 16" counter.

pulse amplitude must be maintained within specified limits. To provide for reliable counting with pulses of undefined origin, a pulse shaper (Schmitt trigger) generally is inserted between the counter input terminal and the first E-J stage.

Because such an E-J counter is not provided with a self-contained means to display the states of the cascaded counter stages, small incandescent lamps can be lit by means of individual drivers, such as shown in Fig. 3-84, to avoid loading the E-J stages. Gas display tubes require about 60v for striking, but by means of a suitable bias, the additional voltage the transistors are to provide to initiate the glow can be reduced to some 20v, a problemless level.

BINARY-CODED DECIMAL COUNTER

With the four truly binary counting stages B1, B2, B4 and B8 in Fig. 4-32, the count proceeds up to 15 and then returns to 0. This operating mode is satisfactory if you are only to watch over the flashing rythmn of the display devices; but if a digital display is required, the decimal numeration system we are accustomed to is imperative. A binary counter can be decimal-coded by injecting a carry pulse back to a preceding stage in a way to clear the count (return to 0) past the binary equivalent of decimal 9. The decimal numerals 0 ... 9 get their binary expression 0000 ... 1001 as formerly, but then, instead of 1010, 1011, etc., it starts again 0000, 0001, 0010, etc.

There are a number of solutions to this problem; the code, said "1 - 2 - 4 - 8," we will discuss now is just one of them. Beside the four E-Js, the circuit in Fig. 4-34 contains one AND

gate which delivers an output only if **both inputs** are present. Up to the eighth pulse, this gate is open and transmits pulses to B2; the eighth pulse tips over B8 and closes the gate. Until now, the pulses delivered by B1 and applied to the 0 input of B8 had no effect on this element because it already was 0. But the ninth pulse tips B1 over to "1," and the tenth pulse, throwing B1 back to 0, tips B8 again, opening the AND gate. This clears the counter and all the E-J are now back to 0. As there are 10 states now, this truly is a **decade**. The output of B8 injects a carry pulse into the tens decade input whose output controls the hundreds decade, and so on if there are any more decades.

The two numerals, 0 and 1 as a binary decade reads are called **binary digits**, or **bits**. The "weight" of a bit is its order, from right to left; the first (right-hand) bit thus has the lowest weight. The weight of the next bit is ten times that of the first one, and so on.

The circuit diagram of a binary decimal counter is shown in Fig. 4-35. The display is made up here of 10 glow lamps numbered from 0 to 9, arranged vertically according to Fig. 4-36. (This is not an inline readout!) There are two rows of odd and even-numbered lamps, fed by means of resistance networks in a way to light only the relevant tube for each counter position. Both tube rows are separately biased by floating supplies. The heavy lines in the schematic indicate feedback paths for clearing at the tenth count.

Such transistorized counters can be made to operate very rapidly: 10 MHz and up. This means that 10,000,000 events can be counted in a second! Decades composed of eight transistors and some diodes and resistors are assembled on plug-in printed circuit boards, and a number of such boards are put together to form a self-contained counter. Mass production of

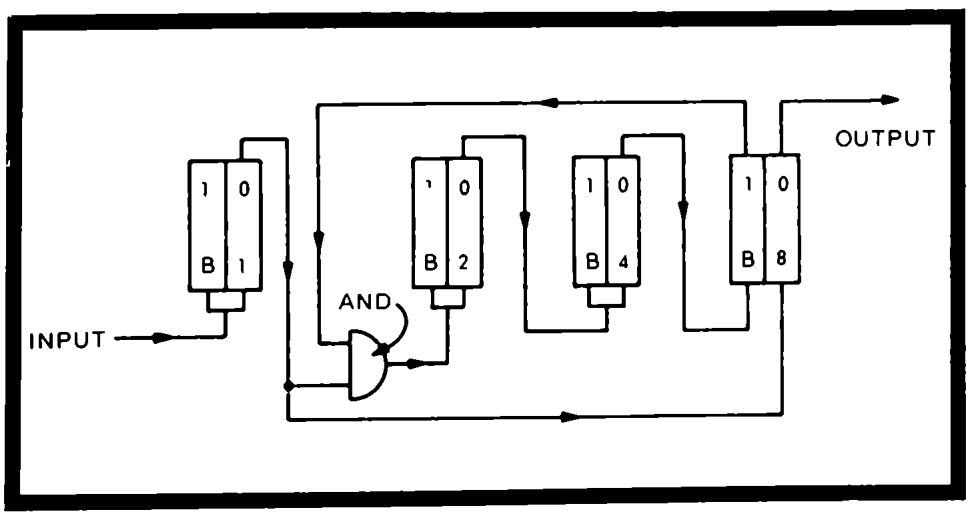


Fig. 4-34. Binary-coded decimal counter diagram. It uses a scale of 16, modified to clear after decimal 9.

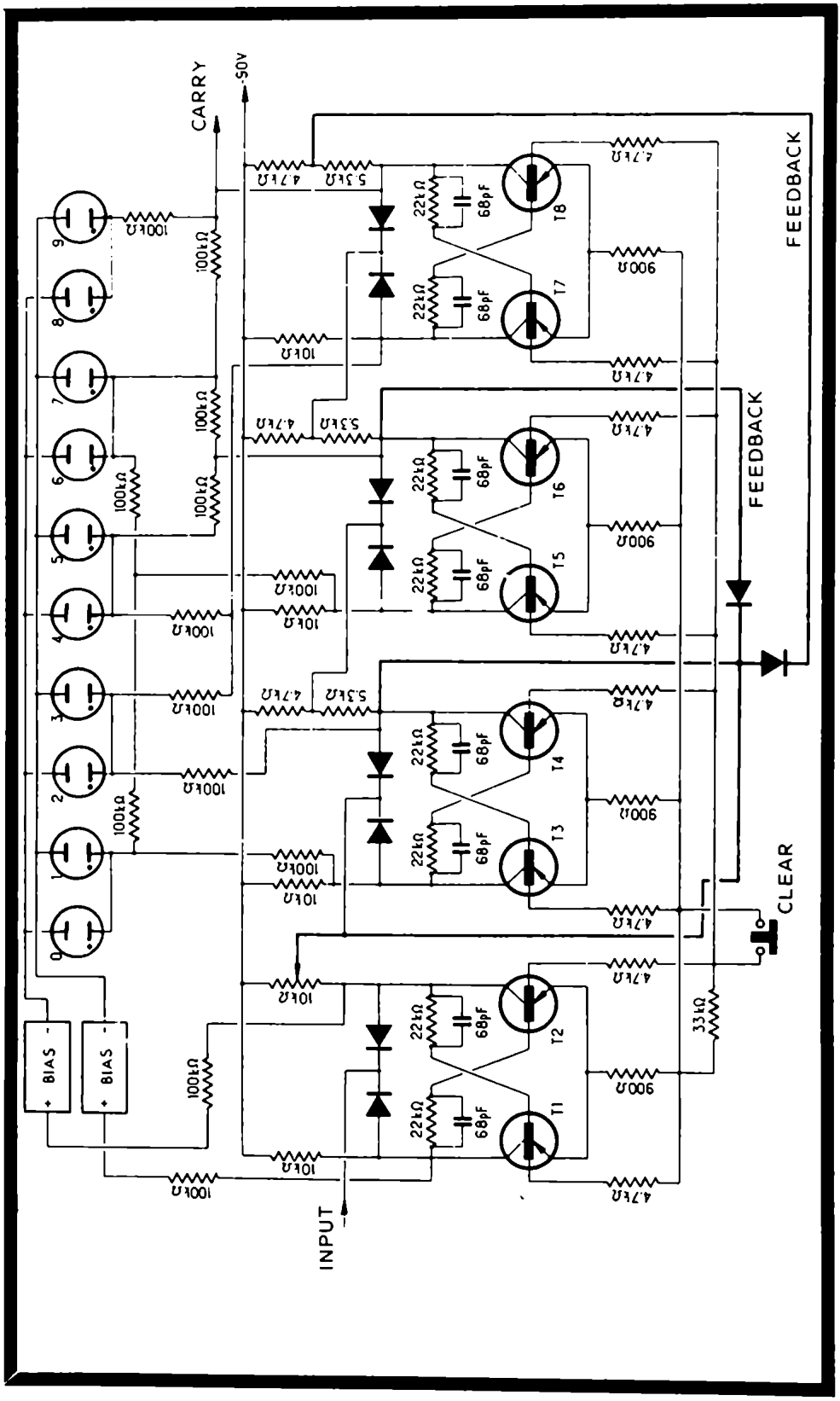


Fig. 4-35. Binary-coded decimal counter with 10 glow-lamp displays.

such printed circuits is economically feasible, using semi-automatic assembly techniques such as dip- or wave-soldering, and components are inexpensive now; but transistor circuits gradually are being superseded by integrated circuits, with a single chip supporting a complete decade. And with MSI and LSI techniques coming, complete multiple decade counters arranged on a single silicon chip will soon be marketed. This is a far cry from that clumsy old Univac computer whose E-J circuits were equipped with vacuum tubes (four duo-triodes per decade).

PRODUCTION LINE COUNTING

For industrial counting, a suitable transducer is needed to detect the fabricated items carried by a conveyor system, and to issue a counting pulse for each detected item. Photoelectric transducers are easy to set up and to use. In the arrangement in Fig. 4-37, the moving items to be counted intercept a light beam transmitted by a light source to the photoelectric detector; every time the light beam is intercepted, a pulse is sent to the counter.

A photoelectric device usually is adapted to the shape and-or brightness of the product to be counted. If it presents reflective flats, such as tin cans do, the light beam can be reflected by the glossy can cover to reach the detector device. A dark conveyor is used to show up clearly the bright items (Fig. 4-38).

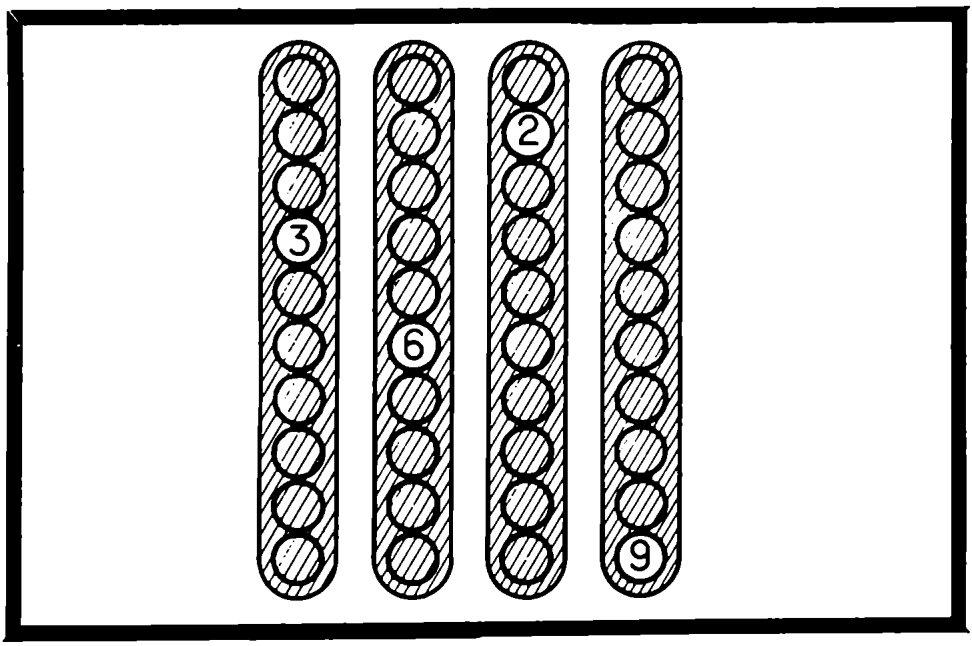


Fig. 4-36. Display of the number 3,629 by vertical rows of glow lamps.

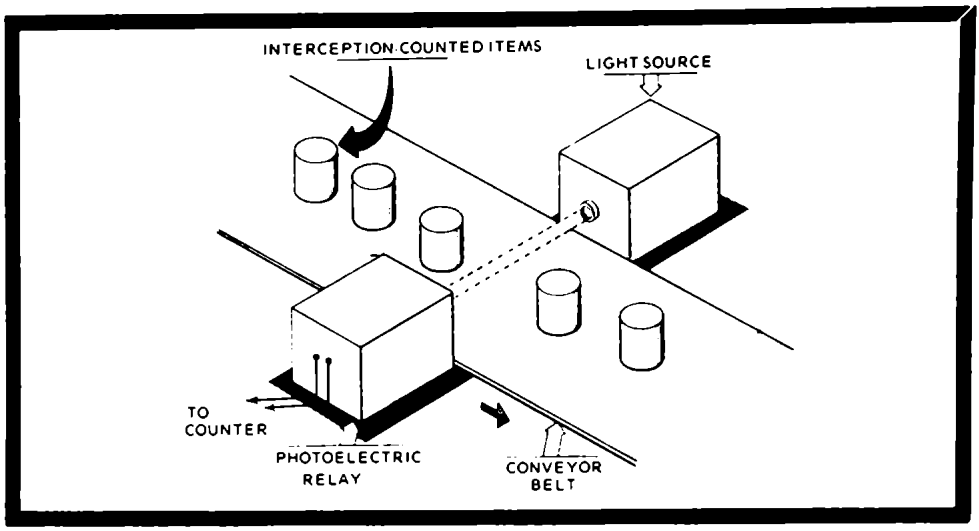


Fig. 4-37. Drawing illustrating the principle of photoelectric counting by light-beam interception.

The drug and pharmaceutical industry must fill large numbers of bottles with an accurate number of tablets or pills. Industrial packing lines use devices such as that shown in Fig. 4-39. The tablets are fed by a vibrating hopper that makes them progress orderly aligned in a groove where they are counted photoelectrically. The number of tablets to be contained by every bottle is preset on the counter. When this number is reached, solenoid S1 is energized and it switches the deflector plate to direct the tablets to the free filling channel, A or B, where another empty bottle is waiting. Tablets stream into the bottle, carried by the conveyor and retained by a turnstile driven by actuator solenoid S2. After a delay necessary for completing the bottle filling operation, S2 is energized and it rotates the turnstile by 90 degrees, setting free the full bottle, which is driven by the conveyor to the bottle packing station. Meanwhile, the same filling operation proceeds on Channel B, whose turnstile is controlled by actuator S3.

There are many tasks of this type that could be achieved—perhaps equally well—by purely mechanical or electromechanical means; but the extreme flexibility of the electronic solution often will be lost. Suppose, for instance, that the bottle content is to be changed from 50 to 80 tablets, a change that frequently occurs in production lines. While it takes only seconds to reprogram an electronic counter, it may be costly and time-consuming to modify a mechanical packaging system. Furthermore, the electronic counter can issue daily production reports, and can even be "interfaced" to provide a computer input for production control or cost accounting purposes.

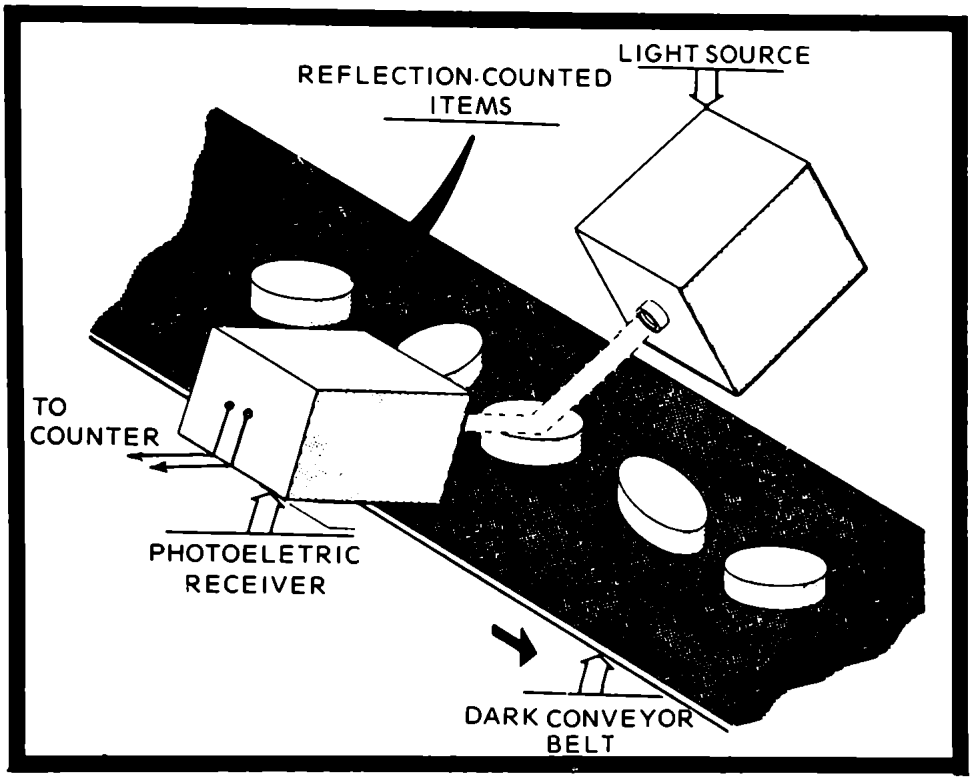


Fig. 4-38. Photoelectric counting systems also use light-beam reflection.

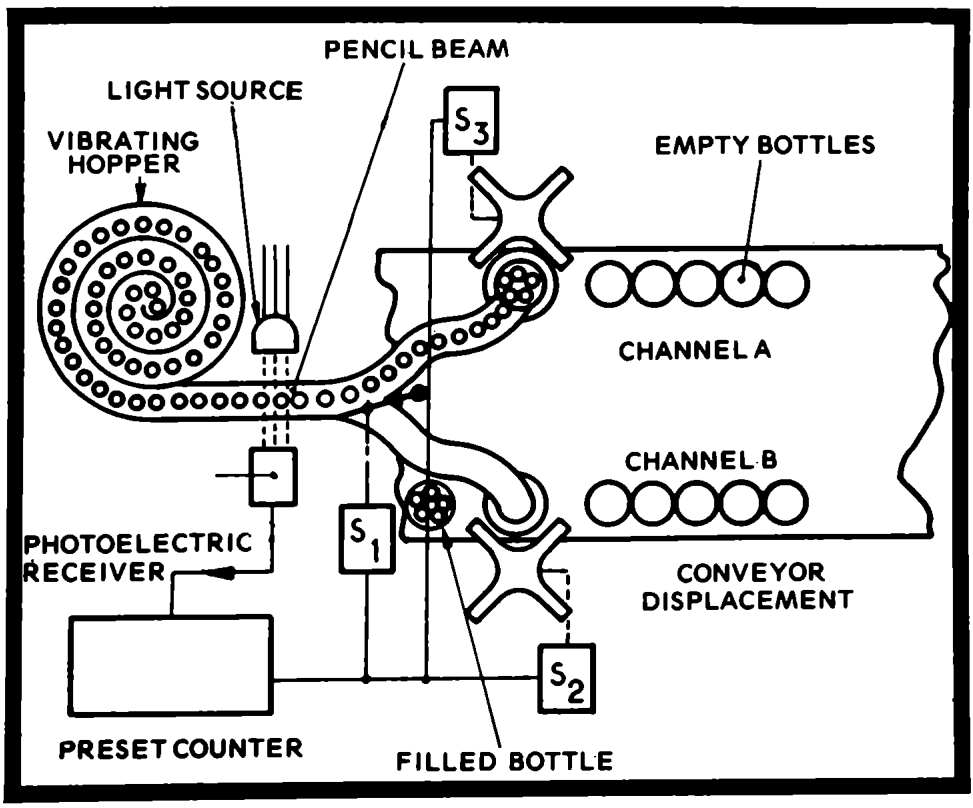


Fig. 4-39. Drawing of an automatic bottle-filling operation.

CHAPTER 5

Digital Process Control

To control the operation of a machine just as a human operator would do, an electronic control system must be able to make **logic decisions** imposed by a given set of conditions. The orders transmitted to the machine under control are subordinated to the presence of some conditions as well as the absence of others. All these conditions are expressed without any ambiguity in binary language by **yes** (1) and **no** (0), corresponding to the conducting and not-conducting states of a switching device (switch, relay contact, diode or transistor). A control system that is capable of distinguishing all relevant situations by the binary digits 0 and 1 is called a **digital** system.

Since intelligence is **the ability to respond quickly and successfully to a new situation** (Webster's New World Dictionary), even a rather low-IQ (intelligence quotient) human operator is remarkable in the manner by which he can readily adopt himself to a new or unforseen situation: he continuously learns by doing. For instance, take the case of grasping items on a moving conveyor, a task so simple that even trained animals can learn to perform. Under visual guidance, the arm and hand muscles will instinctively adapt and coordinate their efforts in terms of the actual position of the object, once perceived.

The automatic system lacks this common instinct. It is unable to cope with any situation for which it has not been programmed. If the objects to be grasped on the conveyor are randomly positioned, a transducer system (possibly photoelectric) and a comparatively elaborate servomechanism will be necessary to direct the pickup device, so it generally is easier to properly line up the items prior to picking them up.

Without being specially trained, an operator would readily notice an uncommon aspect of one of the items he deals with, such as a different color. For its part, the automatic system can, of course, be fitted with a colorimeter to detect possible color changes, but this obviously adds cost and complexity. If this device is not provided, the machine will be unable to

detect this peculiarity. The system designer, therefore, has to think over—step by step—all the operations to be carried out by the machine, consider all situations that possibly could develop, and devise the actions he wants it to take accordingly. Should he overlook a possible situation, the machine would be unable to cope with it, should it develop.

Logic enables the machine to make even the simplest decisions. If your domestic robot is to enter the dining room, it first must learn if the door is open (0) or closed (not open, 1). Suitable transducers are provided to send a signal, 0 or 1, to the robot computer. All digital computers are composed mostly of logic gates that make decisions which result in the application of energy to a motor or solenoid actuator to accomplish the action required by the situation.

The control system instructions are contained in a program which specifies a sequence of operations, such as first open the door (supposing it actually tested closed), and then enter. There are combined instructions in which the decisions depend on the fulfillment of some conditions and on the nonfulfillment of others; for instance, "open the door only if it is closed," means that the action of opening is dependent on the arrival of a transducer signal indicating that the door actually is closed.

Having ordered the robot to open the door, the computer will await the signal, "door opened," to issue the order: "enter the dining room." This is the **feedback message**, "action accomplished," which is the only way the computer can know if the message is free. Since you cannot simultaneously open a door and enter the room, the computer must prevent any (damaging) attempt if doing it. All automatic systems are fitted with such reproting-back transducers. Often there also is a timer to delay the following operation to allow for the previous one to be completed.

The sequence of operations is contained in a **memory**, which instructs the computer of the steps to be taken and of the conditions to be fulfilled to start action. There are electronic and electro-mechanical memories and combinations thereof. The programmer of an automatic domestic washer is a simple electromechanic memory. It is composed of a highly geared-down micromotor that drives very slowly (say, one revolution per hour), a set of rotating wipers which contact fixed contact decks. This provides for a fixed-sequence program where only the duration of certain operations is variable. Industrial production operations can rarely do with such rigid schemes; machine tool programs require quick interchangeability for turning out a different series of workpieces. Perforated cards and tapes or magnetic tape

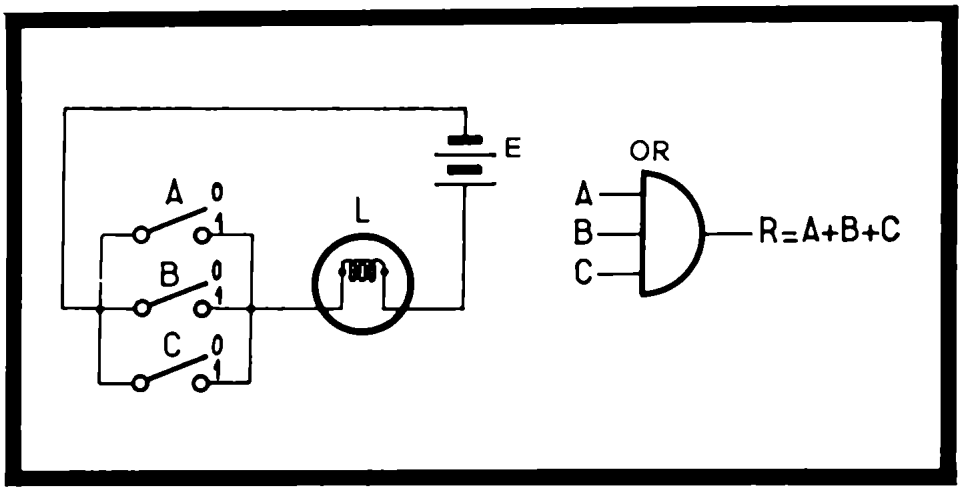


Fig. 5-1. Operating diagram and symbol of the logic OR function.

memories are best suited for programming machine tools, and it is easy to stock such processing information in special processing libraries to permit starting any desired production run of any listed item at a moment's notice.

LOGIC OPERATIONS

The switching functions of a number of 2-state elements (open-closed, on-off...) described by the binary numbers 0 and 1 are treated by an analytical tool developed for this purpose: **Boolean Algebra**, a sort of language of logic. This approach describes the logic operations by the conjunctives AND, OR, NOT, AND NOT and OR NOT.

The logical function, OR, expresses the idea of paralleled switch contacts A, B and C, series connected with a lamp and a voltage source E (Fig. 5-1). To light the lamp, any contact A **or** B or C is to be closed. (If two or more contacts are closed, the result is the same.) The OR operation is symbolized in Boolean algebra by the + sign. Accordingly, the condition for a positive result R (lamp lit) can be written A + B + C equals R; it is understood that A equals 1 means that switch A is closed, while A equals 0 stands for "switch A open."

In Boolean algebra, the OR function is not exclusive as is the conjunction **or** in customary language; it means that closing one of the contacts available is adequate for the action required. Closing more than one contact will bring no change, the lamp will not shine more brightly.

If there are but two variables, there are four equations:

$$0 + 0 \quad 0 \qquad\qquad 0 + 1 \quad 1$$
$$1 + 0 \quad 1 \qquad\qquad 1 + 1 \quad 1$$

It indeed is necessary only to have one contact closed for the result to be 1. Hence, A + 0 equals A (for an open contact paralleling A is inoperative), and A + 1 equals 1 (for if A is shunted by a closed contact, its state is of no concern).

After the Boolean **addition** (there is no subtraction), there is Boolean **multiplication** illustrating **series** connected contacts A, B and C, again in series with lamp L and source E (Fig. 5-2). For the lamp to be lit, it evidently is necessary to simultaneously close A **and** B **and** C. In logic notation, this condition can be written A x B x C equals R, or simply ABC equals R. According to the foregoing, and bearing in mind the physical reality of the circuits, the logic multiplication with two variables can be expressed by the four equations:

0 x 0	0	0 x 1	0
1 x 0	0	1 x 1	1

On the other hand, A x 0 is 0 and A x 1 is 1.

The third basic logic function is called NOT, the complementary or inverting function defined as **the other state** of the (2-state) variable. Thus, the complementary function of a variable A is NOT A, or $\bar{A}$ in logic notation. Hence, $\bar{0}$ equals 1, and $\bar{1}$ equals 0. A double NOT equals a YES, and $\bar{\bar{A}}$ equals A. Electrically, the NOT function is carried out by a normally closed (NC) or break contact relay which opens the lamp supply circuit on the application of a signal to the winding (Fig. 5-3). Electronically, the NOT stage simply is an inverting transistor amplifier stage (Fig. 5-4). With no input signal applied, the transistor is saturated, and its output

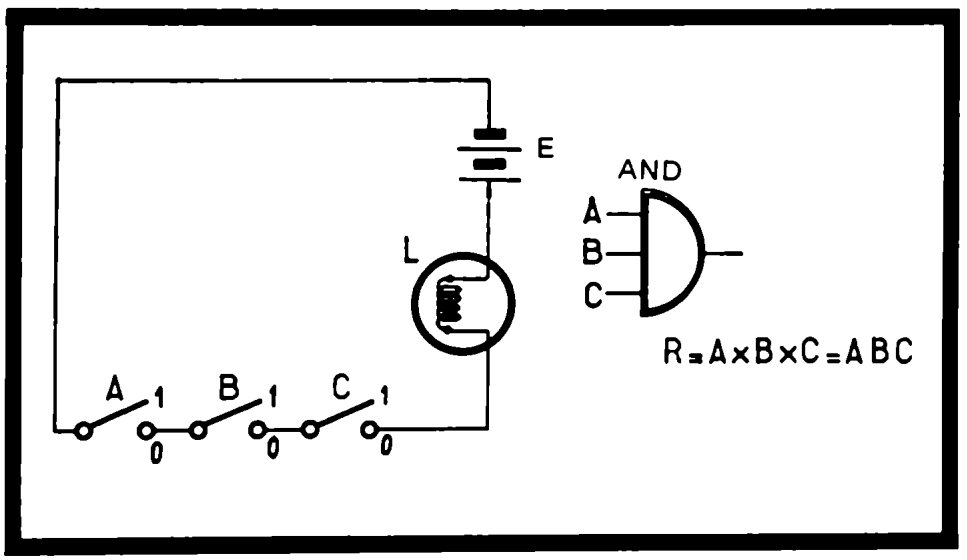


Fig. 5-2. Operating diagram and symbol of the logic AND function.

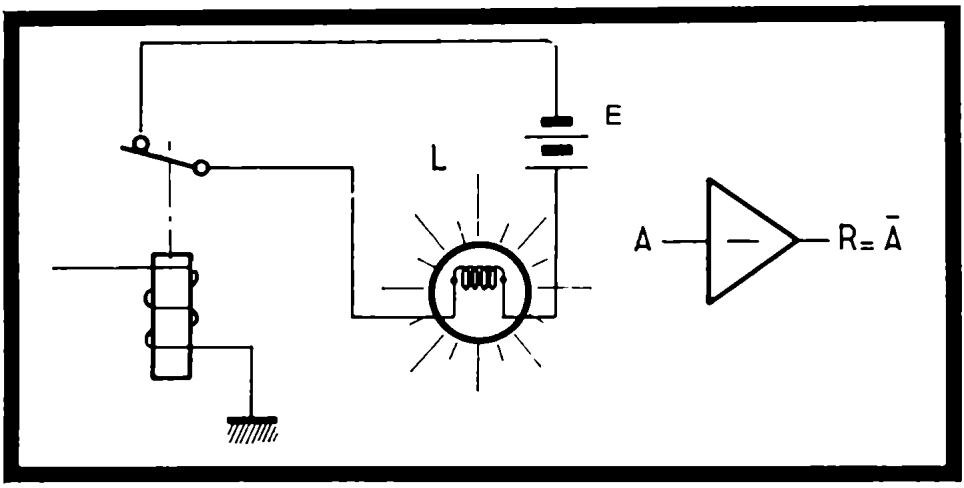


Fig. 5-3. Circuit illustrating the complementary function NOT. With the input in state A, the output is NOT A (Ā).

potential is nearly +E; if a signal with an amplitude of +E is injected at the base, the transistor is blocked and the output voltage is zero. Hence, A is inverted to $\bar{A}$.

If now we parallel (OR function) a contact A and a NC-contact of such a NOT relay to switch L (Fig. 5-5), the lamp will never extinguish, for one of these contacts will close if the other opens. (Beside demonstration purposes, such a circuit evidently is of no use.) Consider the logic notation, $A + \bar{A}$ equals 1. If the two inputs A and B are independent, the circuit becomes a complemented OR "NOT-OR," or NOR for short (Fig. 5-6), whose logic sum is A + B and the inverted output function $\overline{A + B}$. According to a logic algebra theorem, **the complement of a logic sum equals the logic product of the**

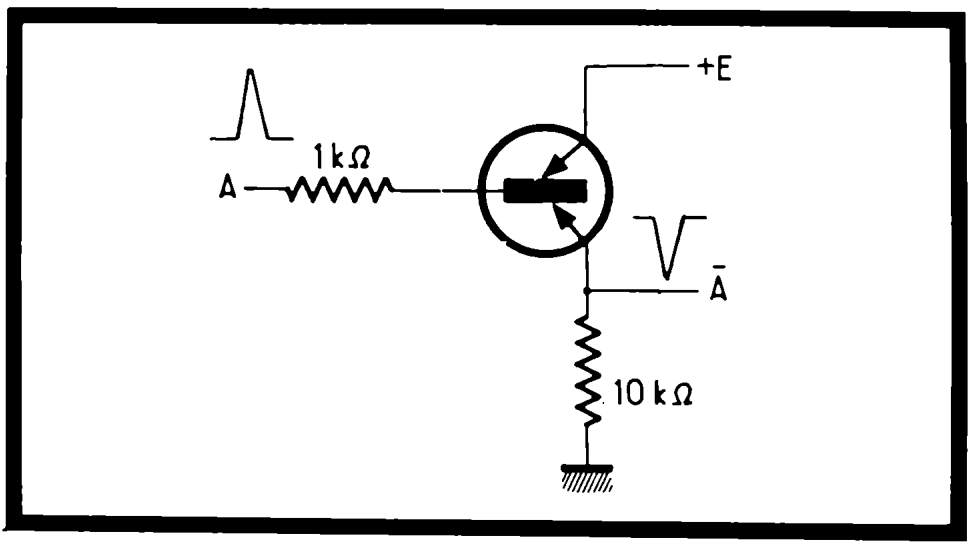


Fig. 5-4. The complementary function is accomplished by a phase inverter.

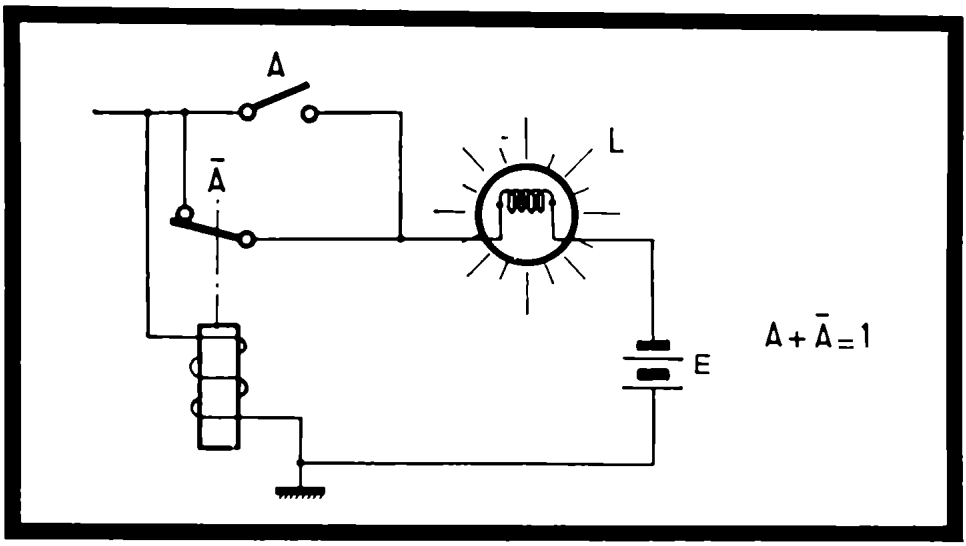


Fig. 5-5. One of the contacts, A and A, in this circuit is always closed, so the lamp cannot extinguish.

complements of the sum elements; hence, $\overline{A + B}$ equals $\bar{A} \times \bar{B}$. If A + B means that at least one of paralleled contacts A and B is closed, $\overline{A + B}$ denotes that at least one contact is open (i.e., possibly both). This is equivalent to two series-connected open contacts $\bar{A}$ and $\bar{B}$.

If we series-connect A with the NC-contact of the NOT relay above (Fig. 5-7), we get A x $\bar{A}$ equals 0. One of the two contacts is always open, and the lamp cannot be lit at all. Though this is another demonstration-only circuit, in the more general case of independent variables A and B we get a **complemented** AND "NOT-AND" contacted to to NAND. With two inputs A and B (Fig. 5-8), the logic product is A x B with an inverse function $\overline{A \times B}$. However, according to the so-called Morgan Theorem, **the complement of a logic product equals the sum of the factor complements**. Hence, $\overline{A \times B}$ equals $\bar{A}$ +

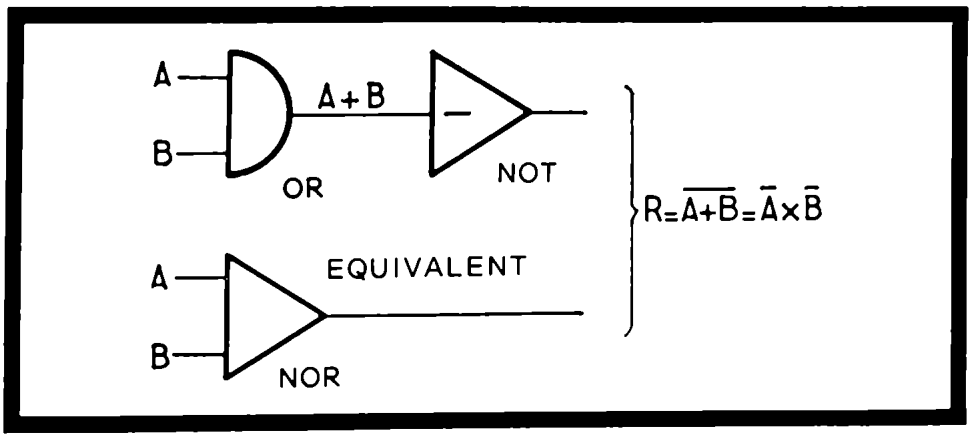


Fig. 5-6. Diagram of the complemented OR or NOR gate.

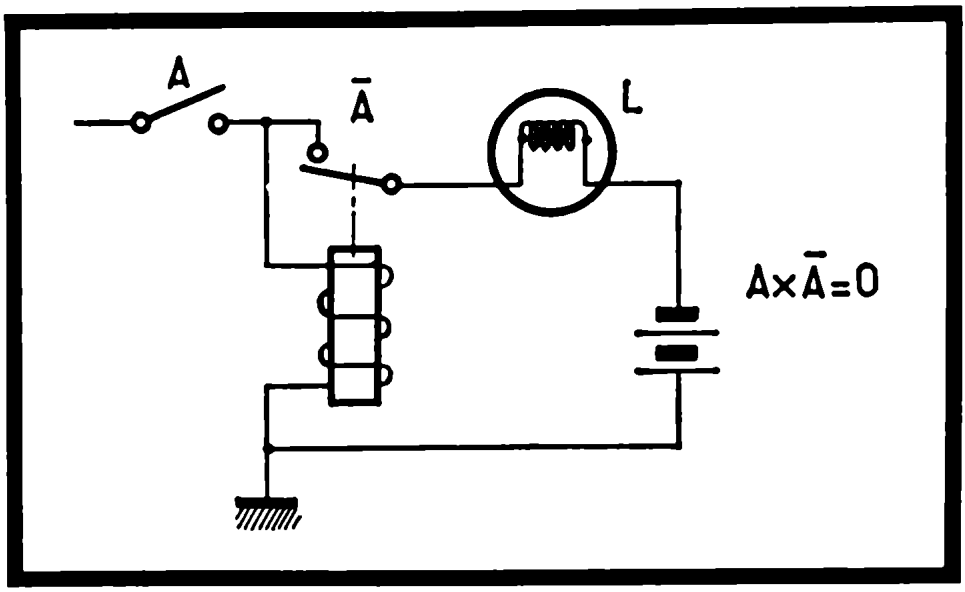


Fig. 5-7. One of the contacts, A and Ā, is always open, so the lamp cannot light.

$\overline{B}$. If A x B describes two series-connected closed contacts, A and B, $\overline{A \times B}$ means that series-connected contacts A and B are open; this equals A and B still open, but now paralleled. (Do not confuse $\overline{A \times B}$, the product complement, with $\overline{A} \times \overline{B}$, the product of the complements, equal to the sum complement $\overline{A + B}$.)

OR, AND, NOT, NOR and NAND are the fundamental logic functions and the action of each should be remembered. When building logic switching systems, however, it is desirable to use the smallest possible number of different circuit types. Furthermore, the production of larger runs of

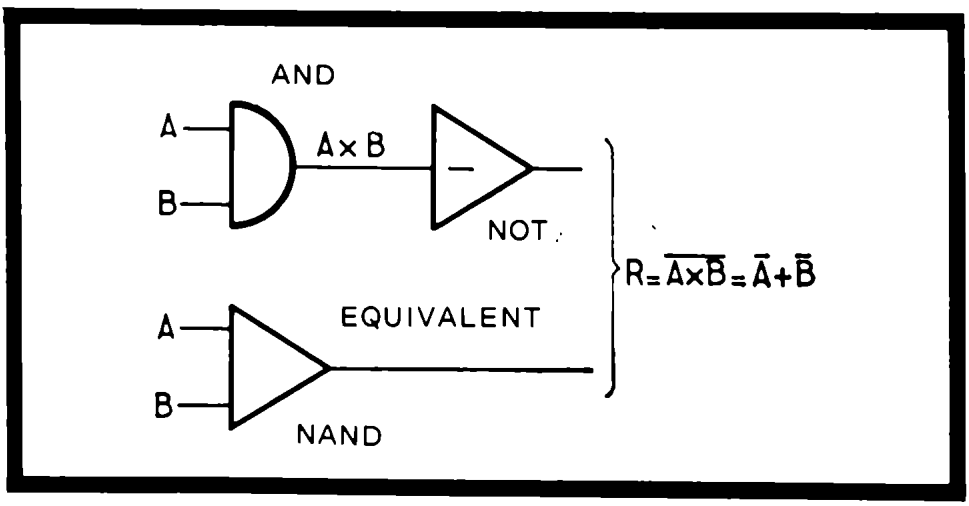


Fig. 5-8. Diagram of a complemented AND or NAND gate.

fewer circuit types is economically attractive. In this perspective, the NOR gate is particularly interesting, for it is possible to set up a complete logic system composed only of discerningly combined NOR gates. NAND gates feature a comparable universality. The apparent complexity of a gate type is of no concern, for in the integrated circuit technology used nowadays, a transistor costs no more than a diode, and enough transistors are available to design reliable circuits in which sophistication is immaterial.

There are some auxiliary logic functions to be mentioned. Signal regeneration actually is a non-reversing amplification that might be termed a YES function. It can be made up of two cascaded NOT gates. Further, there is the MEMORY function. It has been said that a digital signal is a pulse of unspecified tail—short or long duration, it is ephemeral. So a device has to be provided to remember that the signal occured if it did. The Eccles-Jordan circuit in Fig. 4-31 features such a memory, for it can be tipped over by any pulse and it will remain in the new state until it is returned to its initial state by a pulse on another electrode; thus, it retains the information received. The last auxiliary function can be termed TIMING or DELAY; it uses a flip-flop or univibrator (Fig. 3-75, 3-76) "flipped" by the signal, then "flopping back" to its stable state after a preset time, thus forming the desired delay.

TRACING LOGIC DIAGRAMS

The logic functions discussed above allow for tracing the logic pattern of a system, illustrating in some sort of shorthand the working conditions of the system. For instance, consider the operating conditions of an automatic drillpress control having three variables: motor condition M (0 equals stop, 1 equals rotate), cutting fluid flow valve E (0 equals closed, 1 equals open) and workpiece feed W (0 equals no, 1 equals yes). Since each of these three variables can take either of two positions, 0 and 1, there are 2^3 or 8 possible situations described by Table 5-1. The result column indicates whether the particular situation is acceptable or not. Thus, for instance, Situation 6—motor running, workpiece positioned, but no cutting fluid—is not acceptable, while Situation 8—motor running, fluid flowing, workpiece positioned—is acceptable. So the four acceptable conditions can be written: MO FO WO + MO FO W1 + M1 FO WO + M1 F1 W1 equals R. Writing A instead of A1 and replacing AO with its complement $\overline{A}$, we have: $\overline{M}FW + M\overline{F}W + MF\overline{W} + MFW$ equals R.

Fig. 5-9 is the logic diagram of this equation, using four 3-input AND gates corresponding to the four 3-variable ac-

Situation	M	F	W	Result accept	not accept
1	0	0	0	X	
2	0	0	1	X	
3	0	1	0		
4	0	1	1		X
5	1	0	0	X	X
6	1	0	1		
7	1	1	0		X
8	1	1	1	X	X

Table 5-1. Eight possible situations in the drillpress control systems.

ceptable situations. Each AND gate delivers an output only if the three input variables are in the prescribed state. The outputs of the four AND gates are tied to the inputs of an OR gate, featuring an output if at least one input is present. The result R means that the situation is **acceptable** if there is an output, used to enable motor, pump and work feed power supplies.

It is also possible to group the **unacceptable** situations in order to obtain an unacceptable result $\bar{R}$: $\bar{M}F\bar{W} + \bar{M}FW + M\bar{F}W + MF\bar{W}$ equals $\bar{R}$. $\bar{R}$ means that **the drillpress cannot operate if...**, denoting an **interdiction**. Thus, the output (if any) will be used to **disable** the drillpress by interrupting the power supplies, for instance. For security reasons, an interdicting logic system is to be preferred to a permitting one, as is pointed out in Chapter 7.

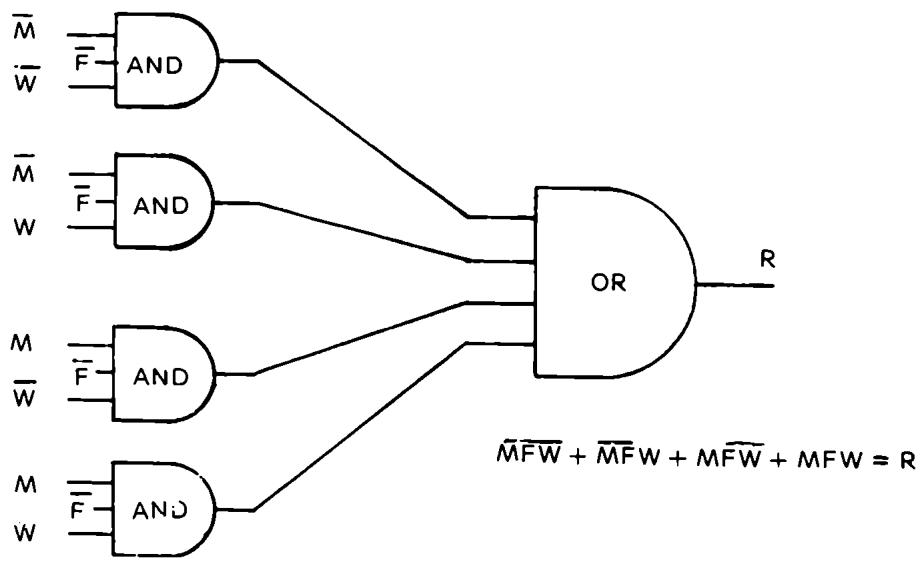


Fig. 5-9. Logic diagram for a drillpress control showing acceptable situations.

The diagram of an interdiction logic system is shown in Fig. 5-10. Now, there are four 3-input OR gates whose outputs are applied to the four inputs of an AND gate. If **all** the OR gates deliver an output, the unique AND gate will provide a disabling signal.

Such logic diagrams are very useful for switching system design by analytical rather than cut-and-try methods. They are also helpful for switching circuit analysis, for they allow a quick check of the current path.

STATIC (SOLID-STATE) SWITCHING

Some years ago, switching functions were accomplished by relays only; but, because of their limited reliability record (contacts can get oxidized and dirty) and life expectancy (moving parts wear out), contactless so-called **static** switching devices are superseding them. Since switches can serve only two possible conditions—conduction and nonconduction—they can be simulated by diodes, transistors and magnetic cores operated as 2-state devices. Though these devices do not present an infinite resistance when nonconductive nor zero resistance when conductive, the current ratio of the two states is rather high and sufficient for all but high-power applications. So (power) relays and contactors are left solely to the final stage; and even there, they can be replaced by thyristors, thyratrons and ignitrons (Chapter 8).

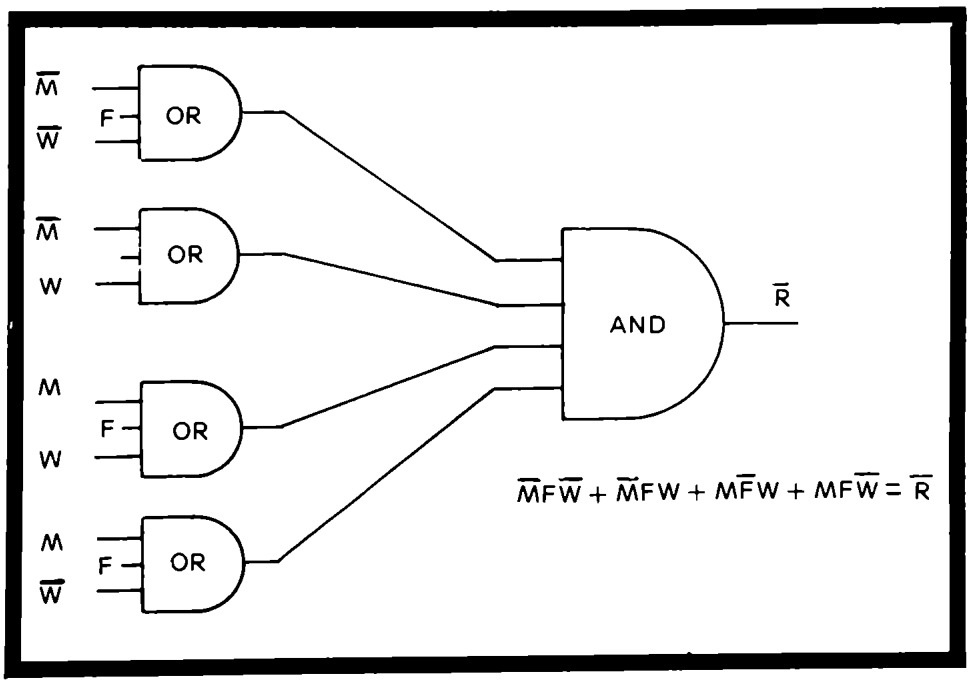


Fig. 5-10. In this drillpress logic diagram, the situations are unacceptable.

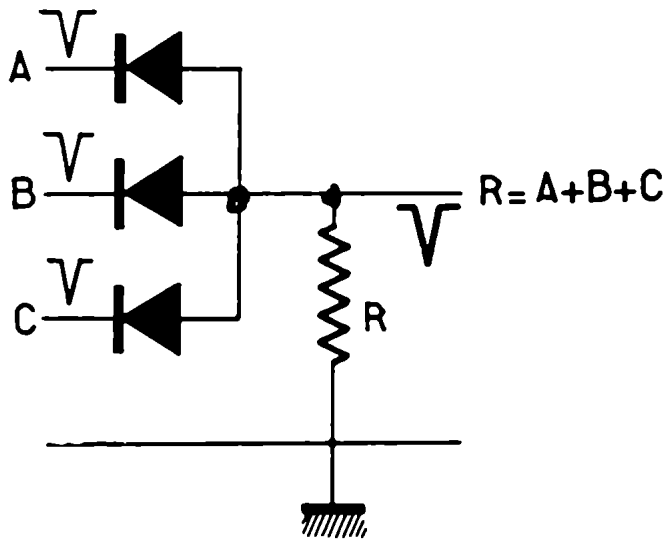


Fig. 5-11. Schematic of a 3-input diode OR gate.

"Static relays" are very compact resin-potted circuits about matchbox size, fitted with pins for plug-in service, or, alternately, printed-circuit boards providing linear printed connectors. Each module is a sealed, interchangeable unit to be discarded should a fault develop. Unfortunately, there is nothing like an industry-wide standard as yet, and circuits of different manufacturers are not interchangeable. These plug-ins are assembled on wired panels that will not be disturbed when an element is inspected. Each element is keyed to prevent the device from being inserted incorrectly.

AND and OR gates can be designed using only diodes. A 3-diode OR gate providing for three inputs, A, B and C, is diagrammed in Fig. 5-11. It is necessary for only one of inputs, A OR B OR C, to be subjected to a negative potential, which biases one of the diodes to conduction, so that the negative potential shows up at the output. In Boolean notation, this means: A + B + C equals R.

In the diode AND gate in Fig. 5-12, the diode cathodes are tied together and returned to the negative busbar through a common resistor (R1). If but one of the diode anodes is grounded, the corresponding diode conducts and grounds the output terminal; hence, there is no output signal. To get an output potential, A AND B AND C must receive negative inputs simultaneouly. The logic operation is A x B x C equals R, or ABC equals R. These two circuit types are often called DL (Diode Logic).

Notice that if positive instead of negative inputs are used, the OR gate in Fig. 5-11 becomes an AND gate, and the AND gate in Fig. 5-12 an OR gate. Hence, to decide on the type of gate you deal with, you have to consider supply voltage and signal polarities.

DL circuits obviously have no built-in gain to make up for any loss, and since gates are often cascaded, it may be necessary to insert amplifiers to restore the signal level. Transistor gates providing the required gain are, therefore, preferred, and in integrated technology, a transistor costs no more than a diode. For these reasons there are (to our best knowledge) no integrated DLs.

The common-collector (CC) transistor configuration, being noninverting, leads to the so-called "positive" logic: AND, OR and YES. The CC amplifier stage in Fig. 5-13 actually is a YES function providing signal regeneration without phase reversal. The AND gate in Fig. 5-14 uses CC-connected PNP transistors. The input pulses (A, B) and output signal (AB) are of the same polarity. Inputs are to be applied simultaneously to A and B to get an output, since both transistors are then blocked. The OR gate in Fig. 5-15 is much like the AND gate, but uses NPN transistors, still CC-connected to retain the noninverting feature. These transistors are normally nonconducting due to a negative bias on the bases in regard to the emitters. Only one positive input on A OR B is necessary to get a (positive) output. These 2-transistor gates are referred to as TTL (Transistor-Transistor Logic).

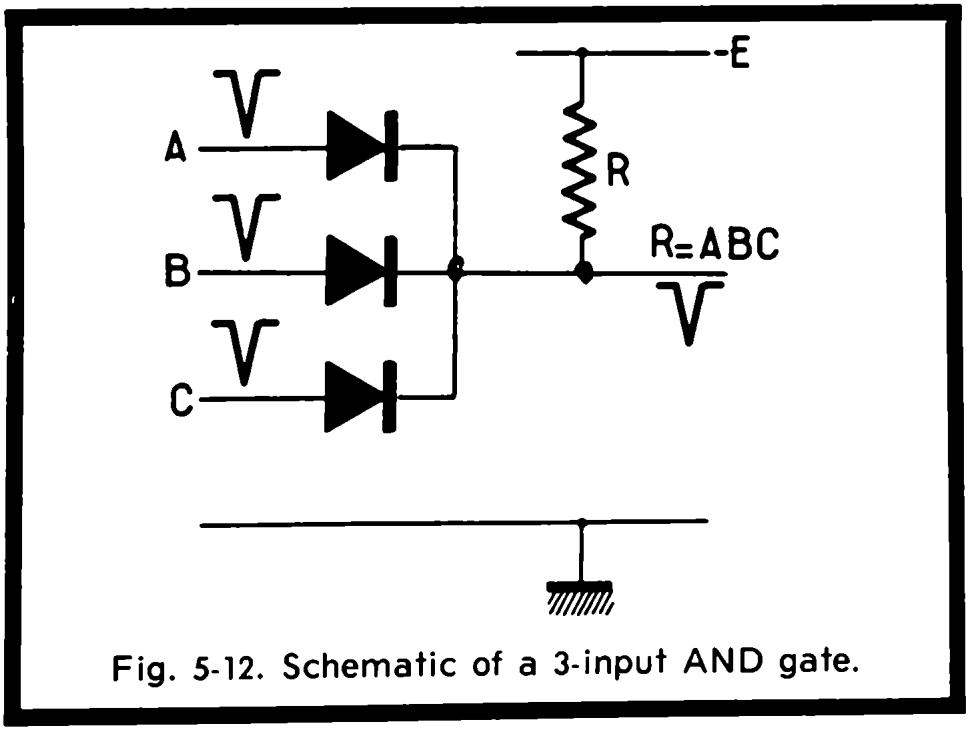


Fig. 5-12. Schematic of a 3-input AND gate.

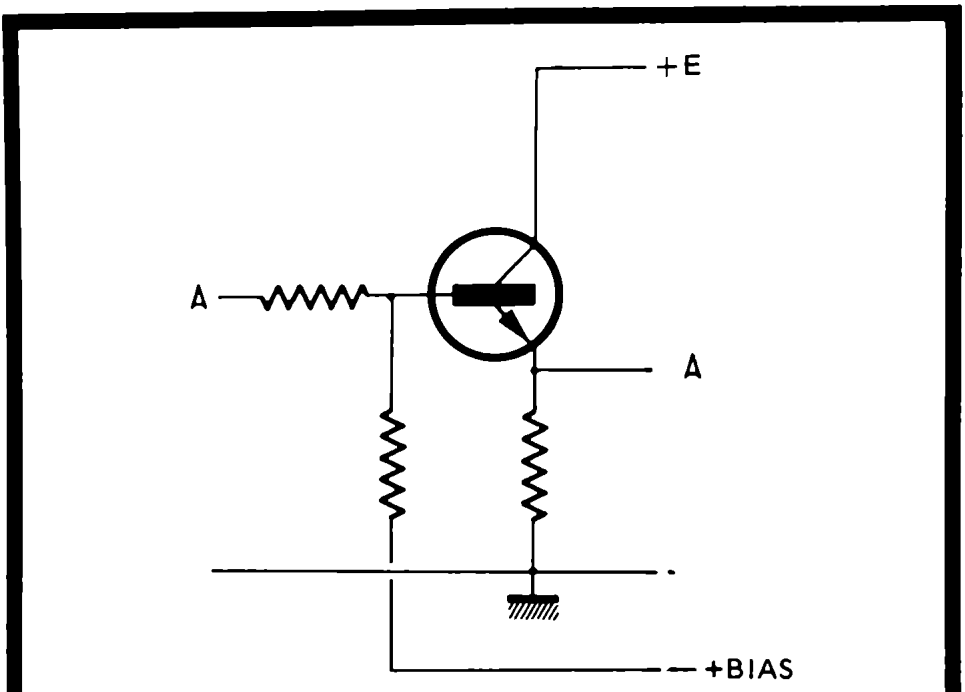


Fig. 5-13. Transistorized YES or regenerative circuit.

Transistors, however, mostly are common-emitter (CE) connected, resulting in a polarity reversal, leading to the so-called "negative" logic functions: NOT, NOR and NAND. Such functions feature a higher degree of flexibility, because cascading two negative functions yields a positive one, while no negative function can be obtained using any number of positive logic circuits. This explains the "universality" of NOR and NAND gates.

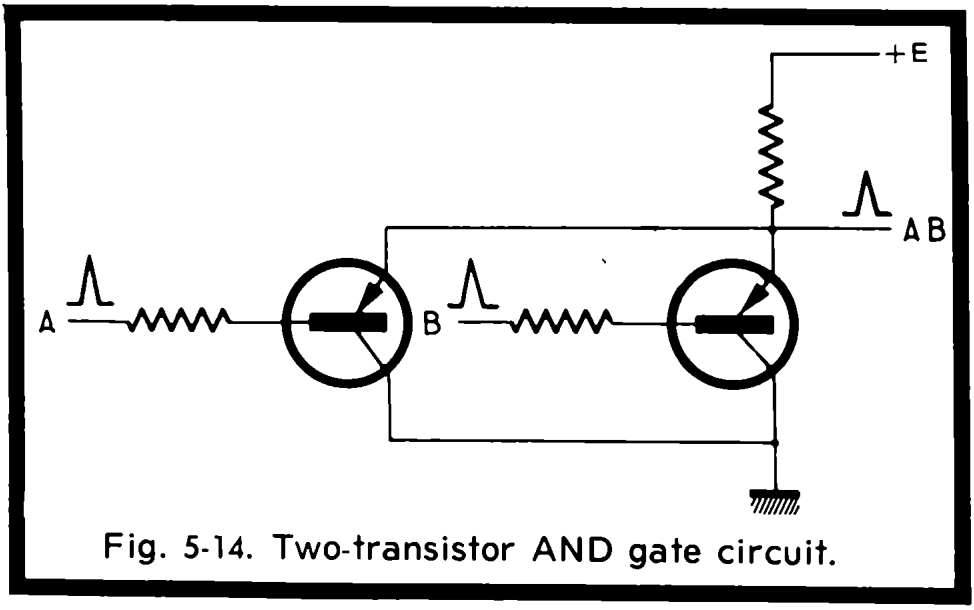


Fig. 5-14. Two-transistor AND gate circuit.

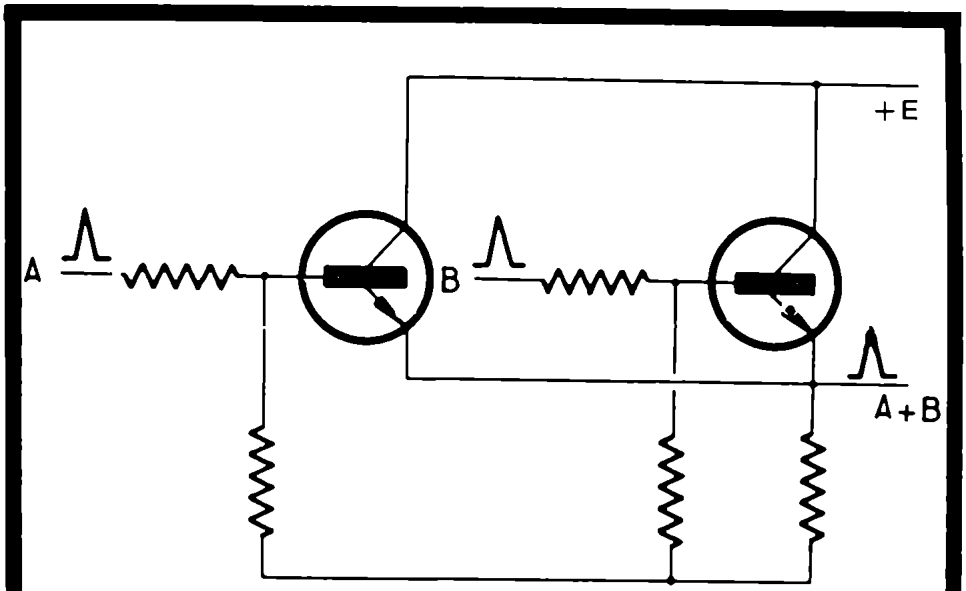


Fig. 5-15. Two-transistor OR gate circuit.

The NOT circuit in Fig. 5-16 is a simple CE inverting amplifier: injecting A, you get $\bar{A}$. Cascading two such circuits, you obtain $\bar{\bar{A}}$ equals A (two cascaded NOTS equal a YES), useful for signal regeneration. A 3-input NOR gate is shown in Fig. 5-17, the inputs are obtained by inserting similar resistors between the base and the input terminals. With no input

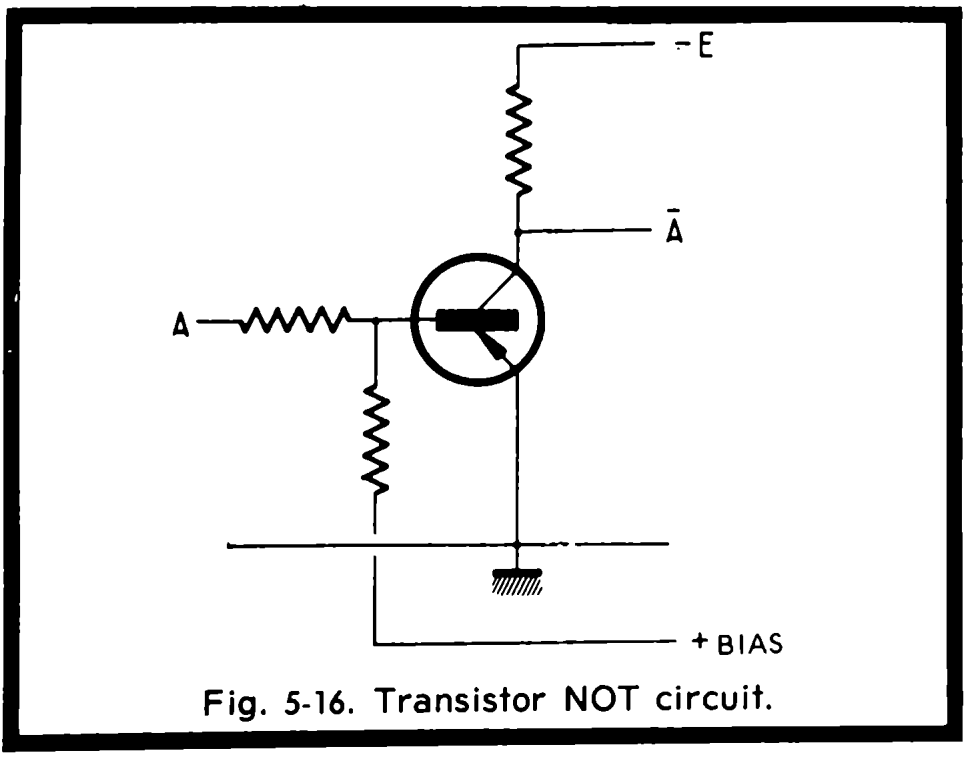


Fig. 5-16. Transistor NOT circuit.

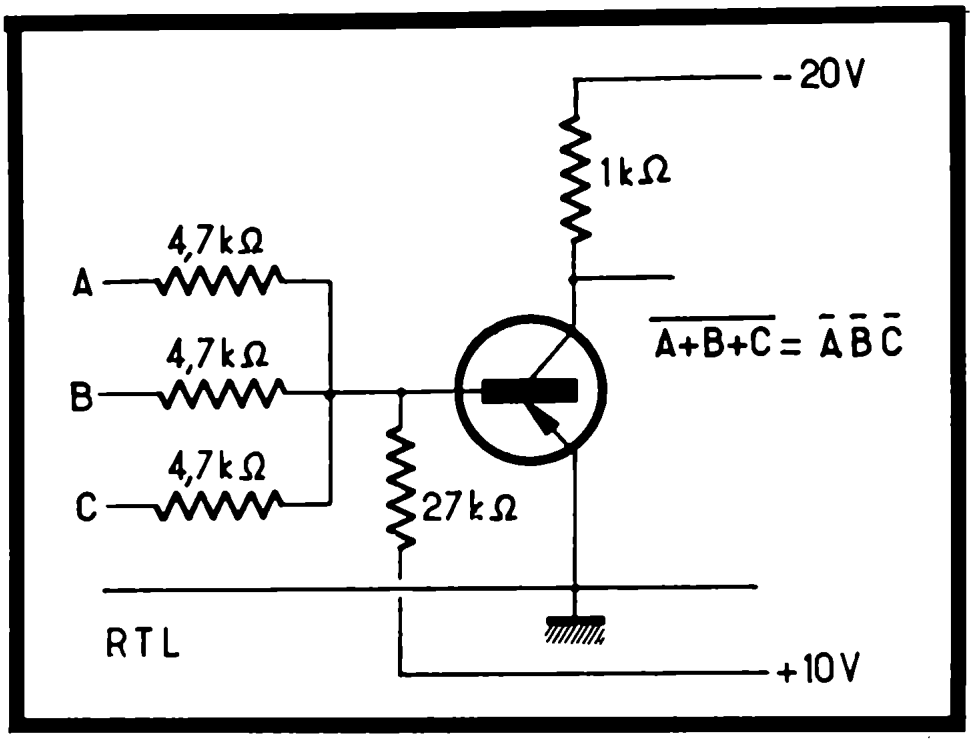


Fig. 5-17. NOR gate circuit that allows all other functions to be carried out.

signal, the transistor is blocked by a positive bias and the output potential is -E. If a negative-polarity signal is applied to A OR B OR C, the transistor becomes conducting, and the output potential falls to a few millivolts. Thus, there is a (negative) output voltage if there is neither A NOR B NOR C. Combining a sufficient number of such NOR gates, the functions AND, OR, NOT and MEMORY can be simulated, allowing for the design any control system with but one gate type.

This circuit is often identified as RTL (Resistor-Transistor Logic). To accelerate the switching interval by reducing the pulse rise time, the resistors are sometimes shunted by small capacitors. (Fig. 5-18.) This circuit, called RCTL (Resistor-Capacitor-Transistor Logic) operates just like the RTL. Replacing the resistors by diodes results in a DTL (Diode-Transistor Logic) circuit in which the behavior is not strictly that of a DL followed by a NOT stage.

As shown in Fig. 5-19A, when transistor T1 is blocked (state 0), Point M is at a potential lower than that of the positive supply bus and diode D1 is biased to nonconduction. Thus, there is a current through R1, R2 and R3. If the T1 base current increases gradually, its collector potential will lower rapidly, and D1 suddenly will become conductive. This abruptly switches the current passing through R1 at Point M

over to D1 and the collector-emitter space of T1 (Fig. 5-19B). Thus, the diode acts as a switch, and there actually is current emission by the input circuit, while in RTL technology, the gate is fed its input current from the external circuitry. This is a NOT-OR, or NOR gate. It is necessary only to change the diode polarity and to adjust the base bias to establish a NOT-AND or NAND gate (Fig. 5-20). In these DTLs, the transistor gain compensates for losses due to diodes and resistors.

Tying together the inputs of a NOR gate results in a NOT circuit. A NOR gate cascaded with another NOT-connected NOR gate equals NOT NOR, or OR. Two cascaded NOT-connected NOR gates yield NOT NOT, or YES (regeneration). The AND function can be simulated by one NOR and two NOT (Fig. 5-21). And cascading two NOR gates with one input of the first connected to the output of the second results in an Eccles-Jordan circuit simulating the MEMORY function (Fig. 5-22).

It can be shown by Boolean algebra that a number of OR gates followed by an AND gate is equivalent to a number of NOR gates followed by another NOR gate. As a practical application, the drillpress control logic in Fig. 5-9 can be worked out using four 3-input NOR gates and a last NOR gate where the four inputs are tied to the outputs of the four preceding NORs (Fig. 5-23).

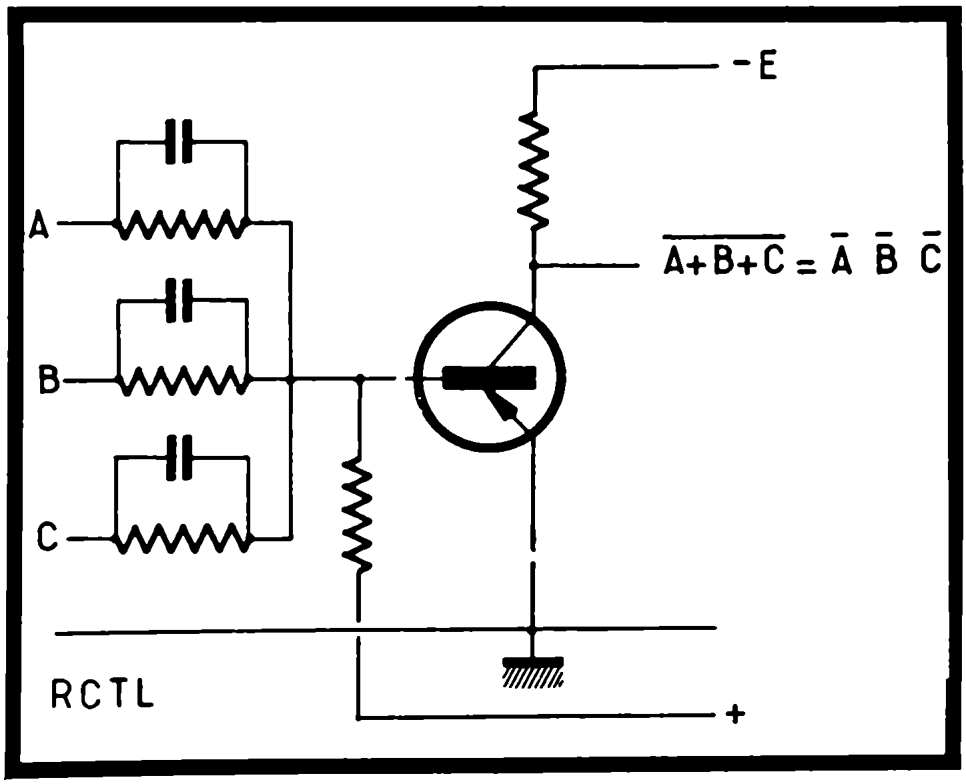


Fig. 5-18. The capacitors across the input resistors adapt this NOR gate for high-speed logic (RCTL).

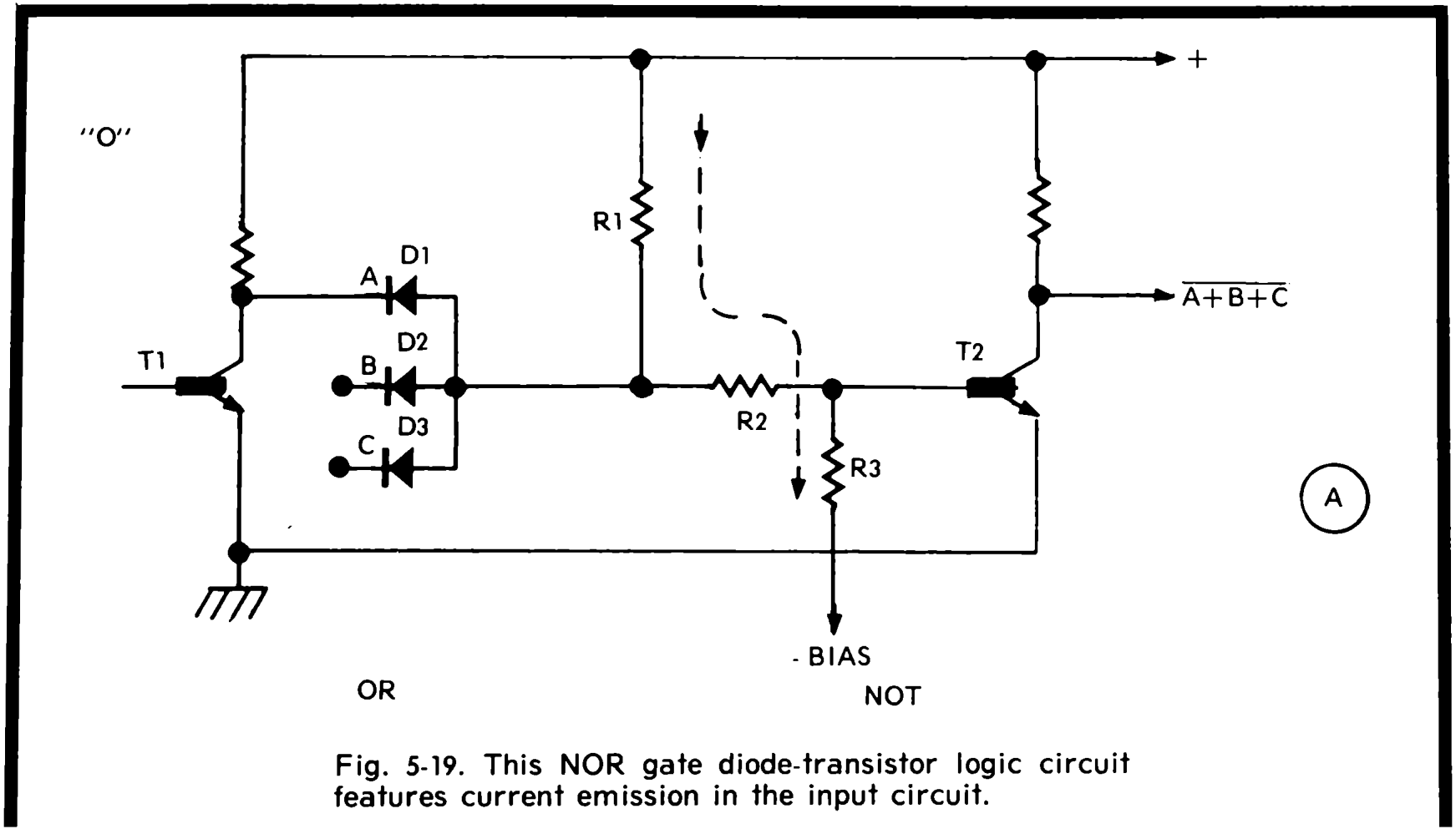
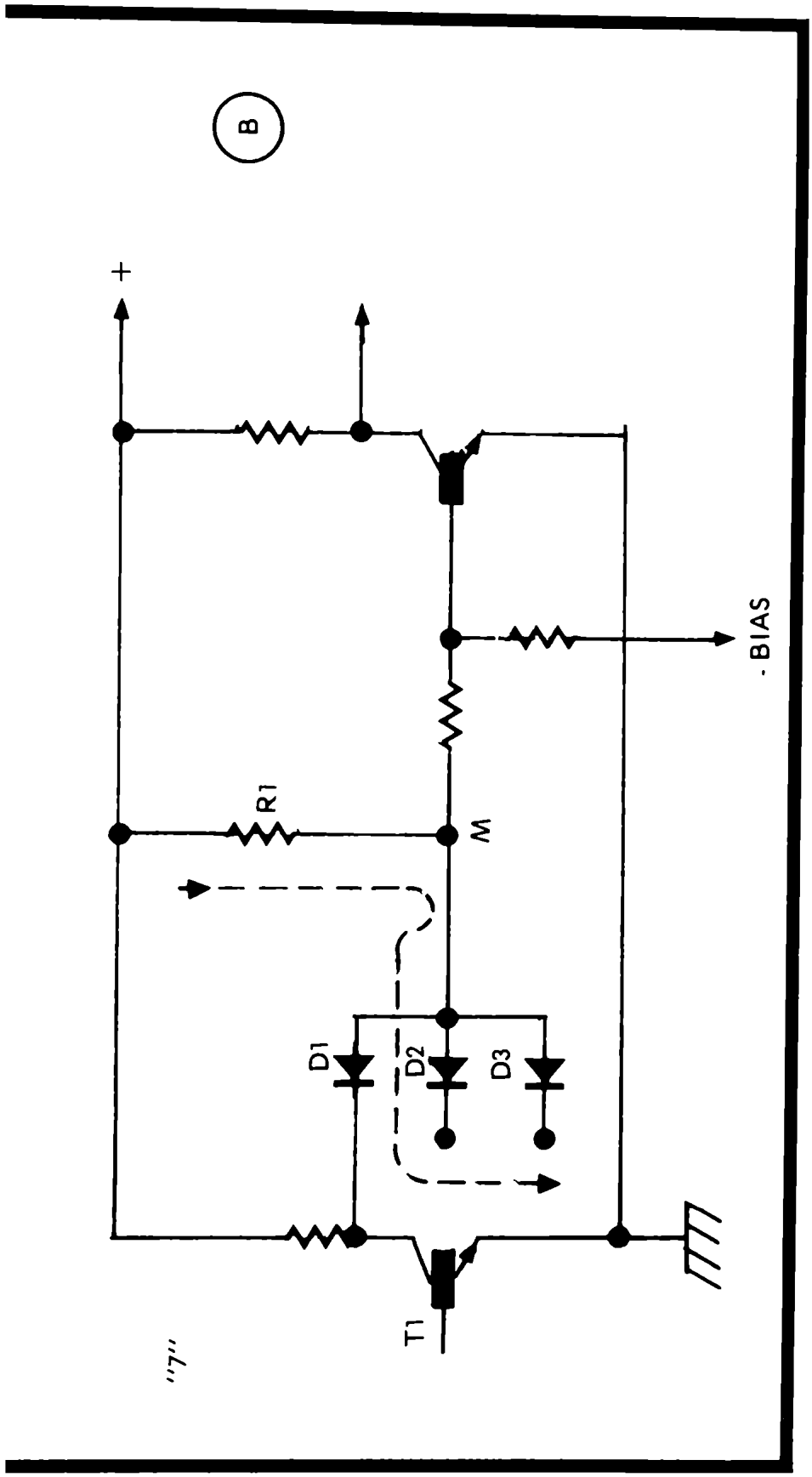
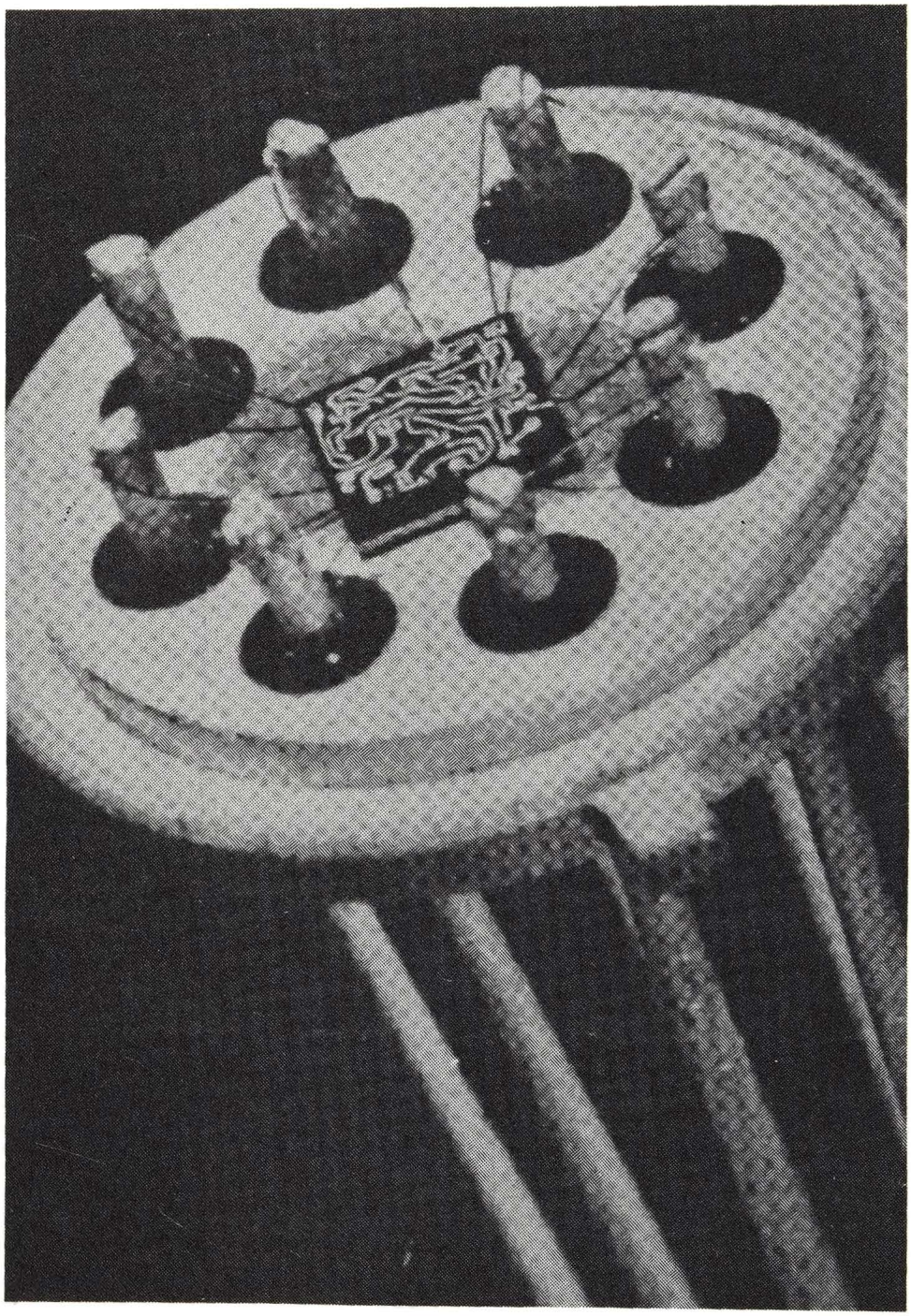


Fig. 5-19. This NOR gate diode-transistor logic circuit features current emission in the input circuit.


B
+
- BIAS
R1
M
D1
D2
D3
T1
"7"



DTL logic integrated circuit (enlarged).

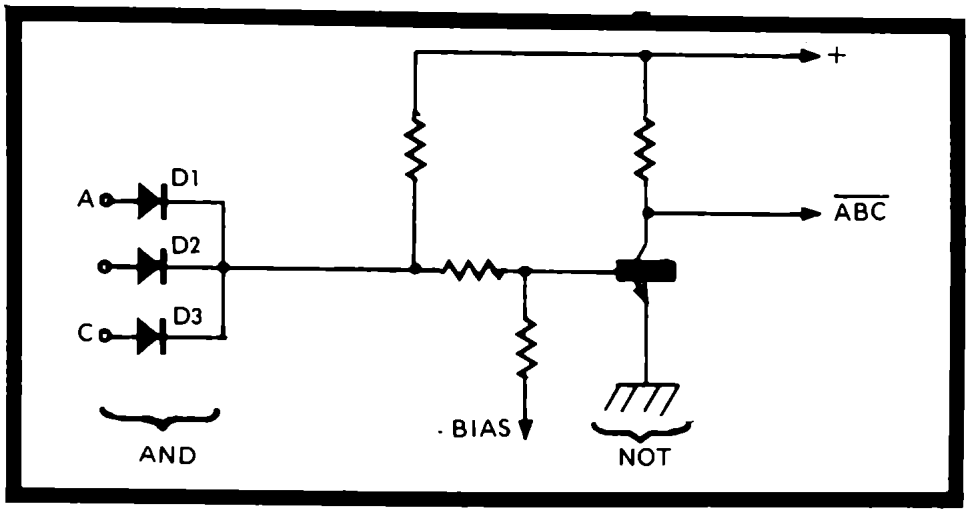


Fig. 5-20. This NAND gate is as universal in application features as a NOR; compare with Fig. 5-19.

DIGITAL INTEGRATED CIRCUITS

In digital electronics, discrete component logic circuits have met rough competition by digital integrated circuits which can be turned out at surprisingly low prices (competition aiding) if mass-produced, and this is made possible by using a unique circuit type for all functions. In addition, there simply are no discrete component assembly and soldering operations that offer reliability and economic production.

Some manufacturers (SESCOSEM) have made the NAND gate, based on Fig. 5-20 their "universal" circuit. This technique allows for an enormous component density: a one-square millimeter silicon chip is said to contain 16 transistors, 4 diodes and 16 resistors! A microscope is necessary to discern

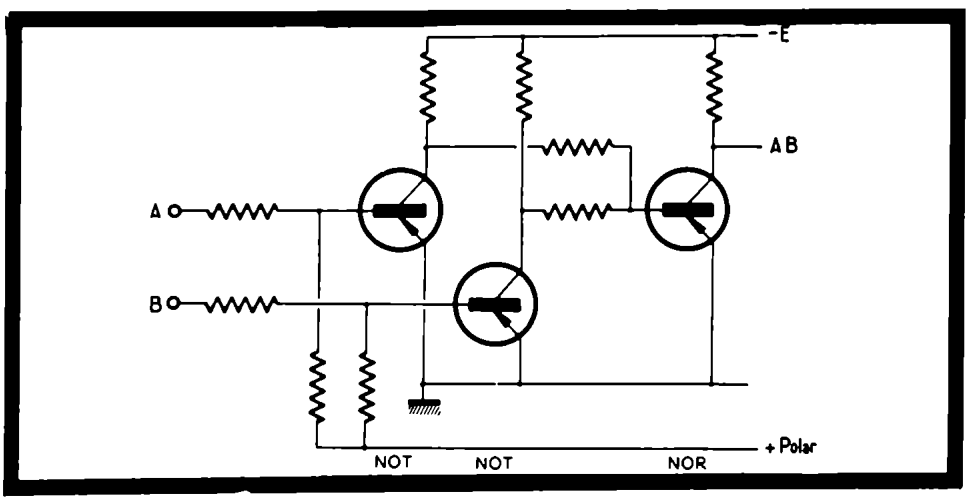


Fig. 5-21. AND function circuit using two NOT and one NOR gate.

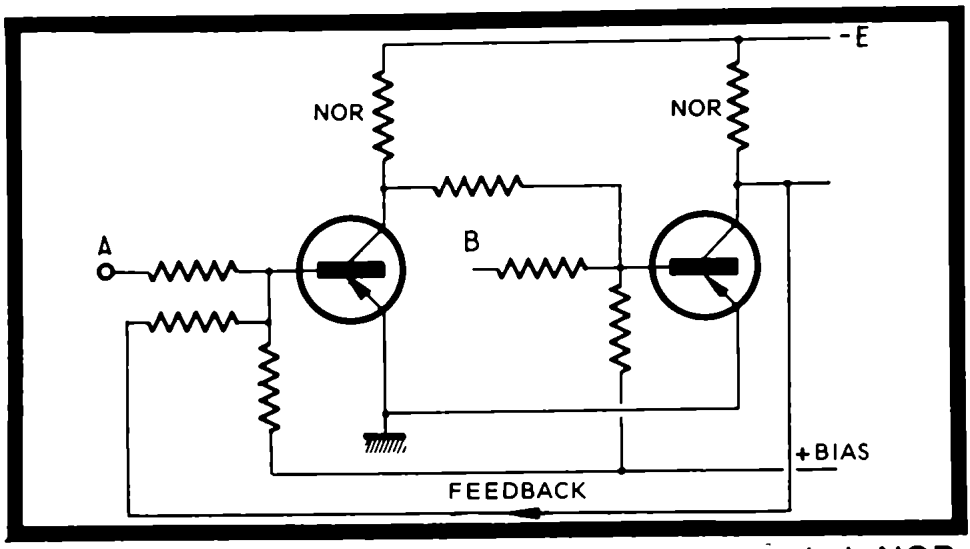


Fig. 5-22. Memory function using two cascaded NOR circuits.

these components. Circuit complexity actually is limited by the number of terminal strips available, 14 or 16 according to the accepted standards for flat rectangular circuits. Thus one device may house:

Four 2-input NAND gates, or
Three 3-input NAND gates, or
Two 4-input NAND gates, or
One 8-input NAND gate.

Since turning out the integrated circuits is the task of the manufacturer, the circuits are represented in the application notes available to customers as closed-contour symbols of conventional shape with referenced inputs and outputs. Using symbolic language, the data tells what happens for the various possible input conditions, but the circuit used to obtain this result is only outlined. Since these circuits are increasingly sophisticated, often only an approximative working principle is stated. On the other hand, the user is no more concerned with component assembly to set up a circuit—he simply puts together logic elements and connects them according to instructions to get a specified result. These circuits offer very high reliability. But should one of them develop a fault, it has to be discarded, since repair is not economically feasible.

A 3-input NAND gate with applied voltage "presences" (1) or "absences" (0) is represented in Fig. 5-24. With the circuits fed from a +15v bus, an input is at logic level 0 if it is grounded or draws some current (from 1.2 to 1.6 ma); its logic level is 1 if it is left free or connected to a potential of +2 to +5v. The output actually emits (1) or draws (0) current,

which allows the circuit to be coupled to others of the same type without any amplifier stage.

The "truth table" below summarizes what happens at the output for the 2^3 or 8 possible input situations:

A	B	C	OUT
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

This table can be sumed up by saying that the output can be 0 only if A AND B AND C are 1.

Leaving free all but one of the inputs of a NAND gate results in the inverting function NOT. The AND function is obtained by cascading a first NAND gate with a second NOT-connected one (Fig. 5-25), and the function OR uses three NAND, two of them being NOT-connected (Fig. 5-26). If this assembly is followed by a fourth NAND, NOT-connected, a NOR gate results. This illustrates the versatility of the NAND gate. It often takes only one chip comprising from one to four NAND gates.

In addition to the basic logic functions, the NAND gate also allows the so-called **auxiliary functions** to be simulated.

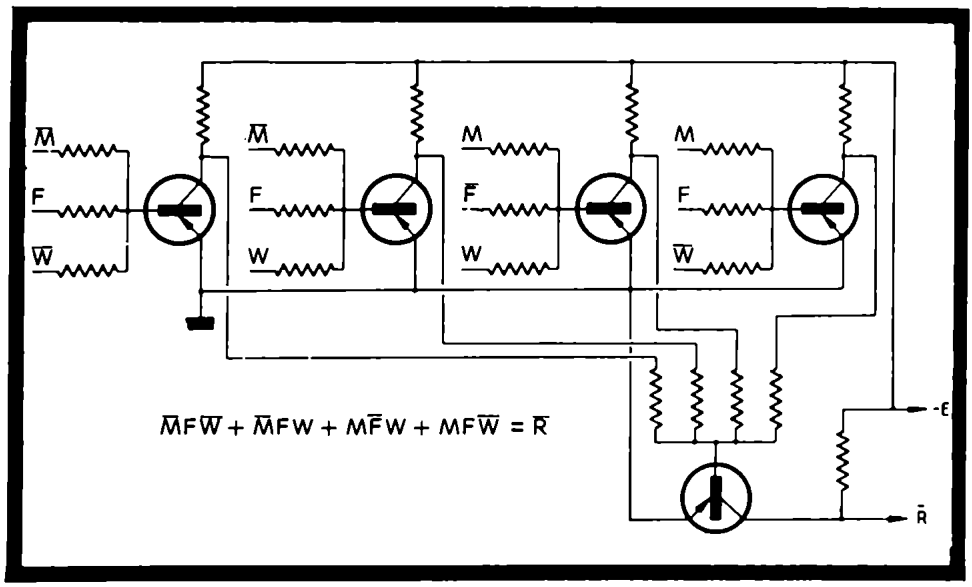


Fig. 5-23. Drillpress logic circuit using only NOR gates.

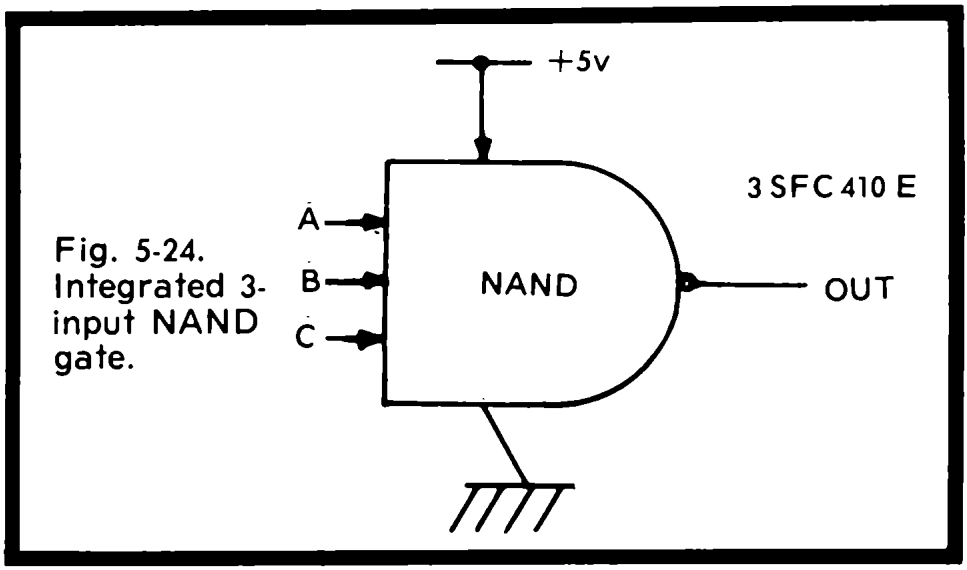


Fig. 5-24. Integrated 3-input NAND gate.

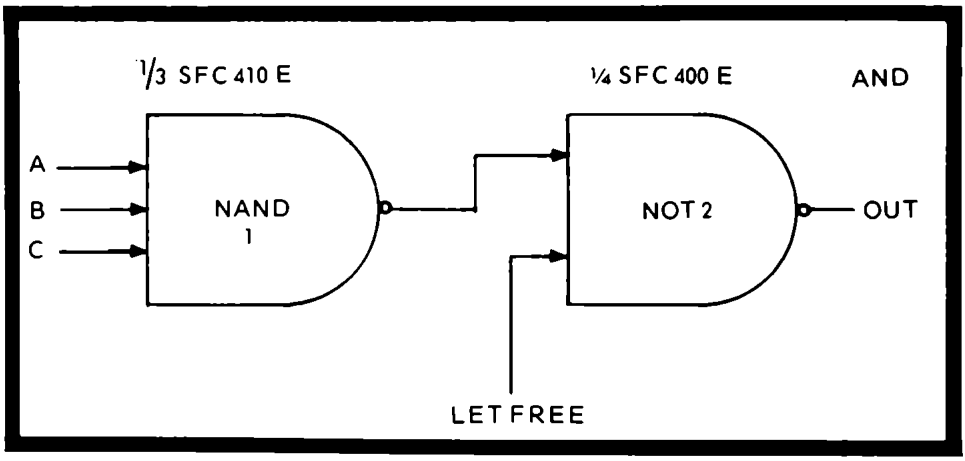


Fig. 5-25. Cascading NAND and NOT gates results in an AND function.

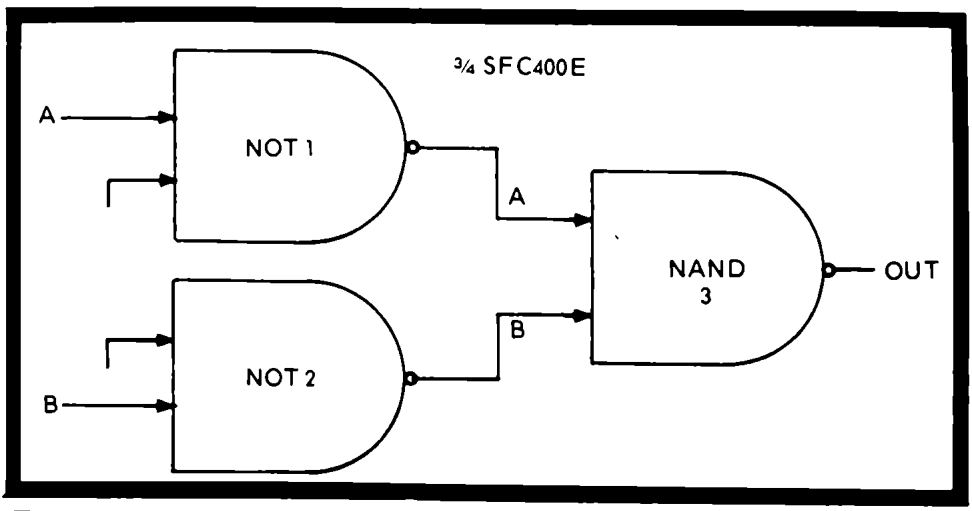


Fig. 5-26. Two NOT and one NAND gate result in an OR function.

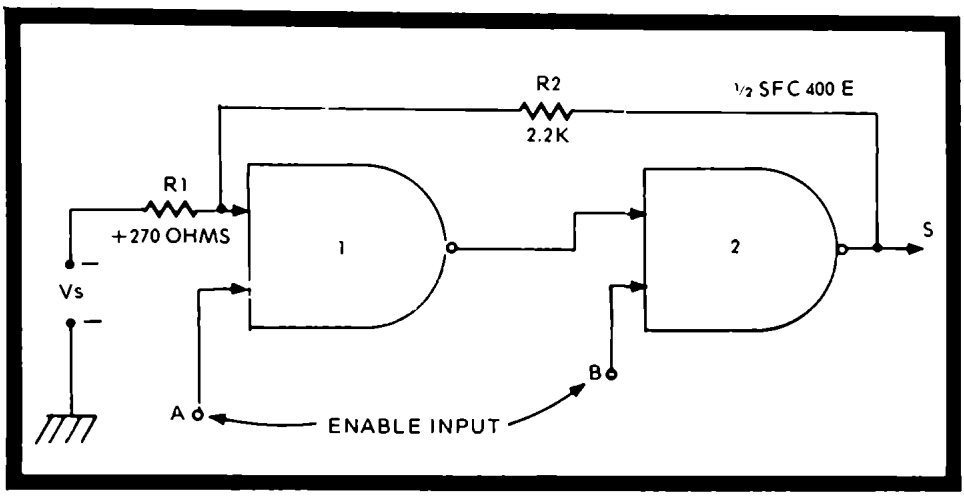


Fig. 5-27. Schmitt trigger circuit made up of two NAND gates.

With two 2-input NAND gates, a **Schmitt trigger** is obtained (Fig. 5-27), useful for waveshaping purposes. The switching which produces the rectangular wave occurs at the 1.3v level on the rising and the 1.0v level on the falling portion of the waveform; the output simultaneously varies from 0 to 3v and from 3v to 0.

A delay-generating **flip-flop** also can be put together, using two 2-input NAND gates (Fig. 5-28). With the numerical values indicated, RC introduces a delay of about 600 ns. Even a **binary counter** can be built with the same NAND gates (Fig. 5-29). Its clock frequency may be as high as 50 MHz. This counter is triggered by negative-polarity 0.2 to 5v pulses with a rise time of less than 20 ns. Since the potentials of Points A and

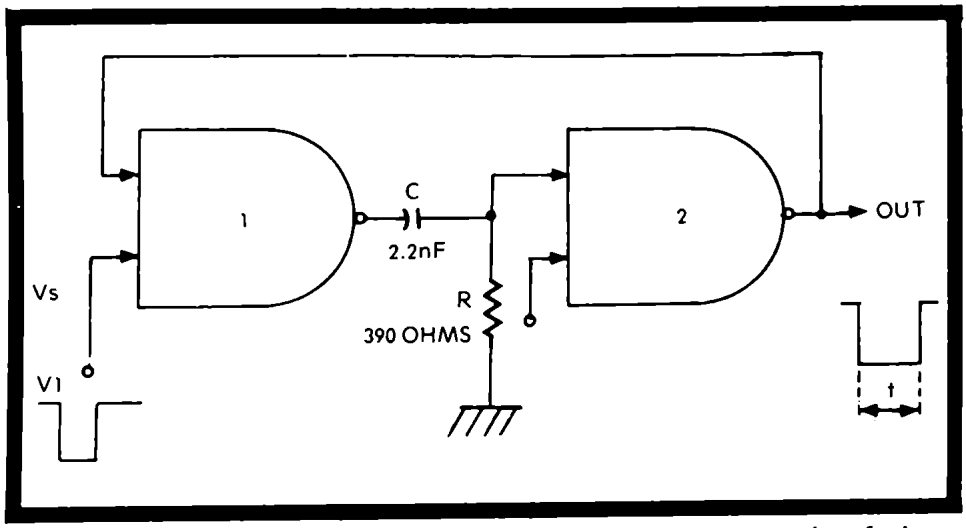


Fig. 5-28. Delay-generating flip-flop composed of two NAND gates.

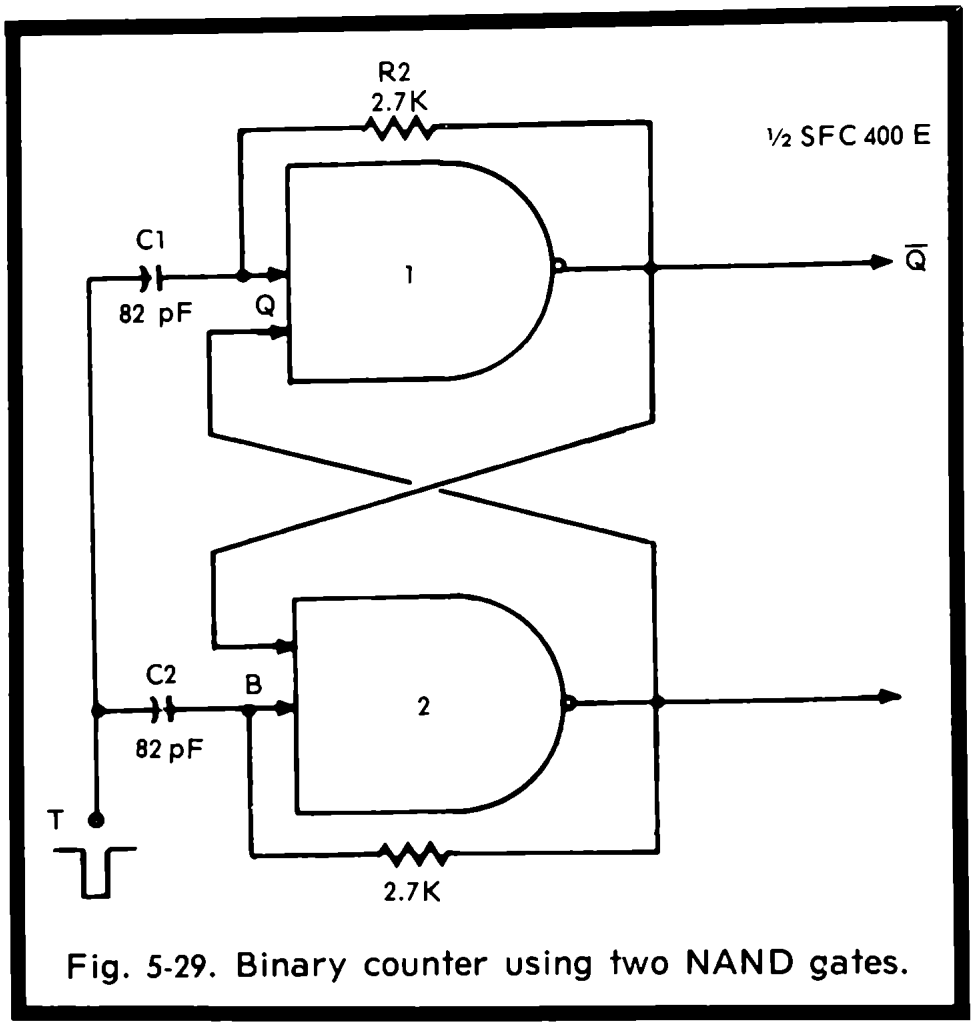


Fig. 5-29. Binary counter using two NAND gates.

B are different (1.5 and 3v), capacitors C1 and C2 are charged to different voltages, and on the arrival of the counting pulse, the lowest potential input switches from 1 to 0; thus, there is a tip over.

The present counter is set up by connecting the NAND gates contained in an integrated circuit; this operation is also directly carried out by the manufacturer who markets several counter types in the round TO116 and flat MP117 casings that allow for 14 and 16 connections. An integrated circuit generally contains two counter stages of medium complexity. The scale-of-two, called "J-K," is attractive due to the fact that a counting pulse simultaneously applied to both inputs J and K tied together surely triggers the tip over. Such a J-K of the "master-slave" type is schematically represented in Fig. 5-30. This is a simple binary counter, the counting pulse is applied to $\overline{T}$ and inputs J and K remain free. There are two output Q and $\overline{Q}$. Q can be tied to input $\overline{T}$ of another J-K to set up a 2-stage binary counter. Grounding Point $\overline{R}$ by closing contact Ko "clears" the counter, setting it back to 0.

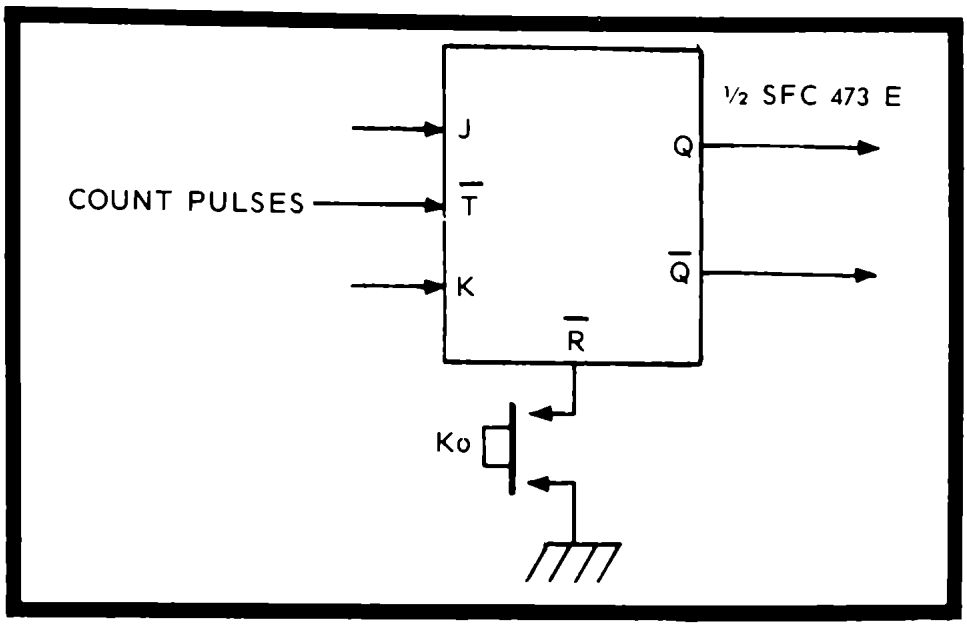


Fig. 5-30. Connection diagram of a J-K scale-of-two circuit.

The counter can operate only if J equals K equals 1; this is obtained by letting these inputs free. J and K can be connected to an input "count enable," which permits counter operation only if these inputs are not grounded (Fig. 5-31). Construction of the interval generator in Fig. 4-27 should prove a lot easier using these integrateds.

The **preset binary counter** in Fig. 5-32 uses five standard integrated circuits. The eight scales-of-two, B1 to B8, allow for

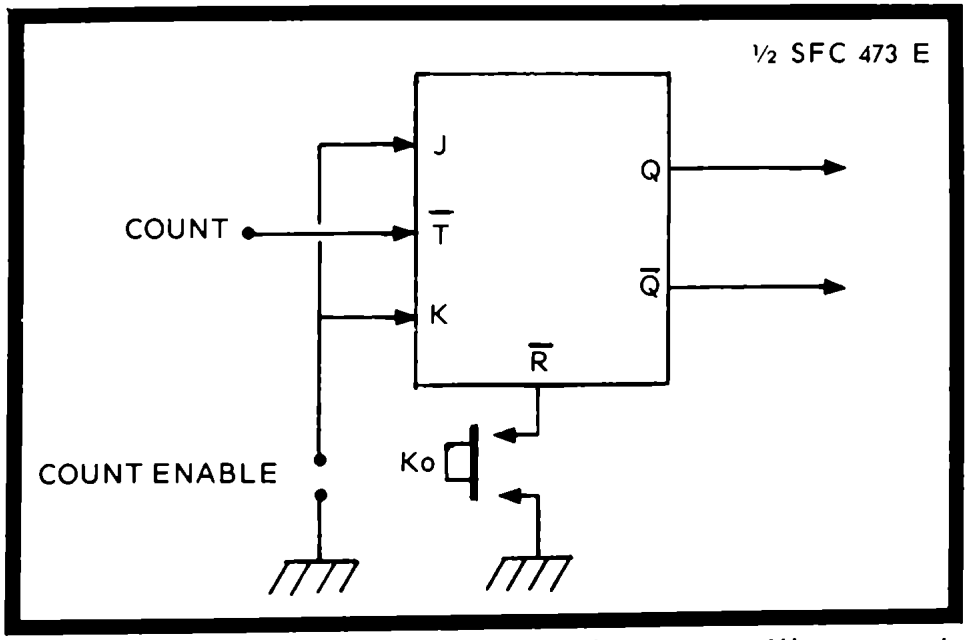


Fig. 5-31. J-K counter connection diagram with a count-enable input.

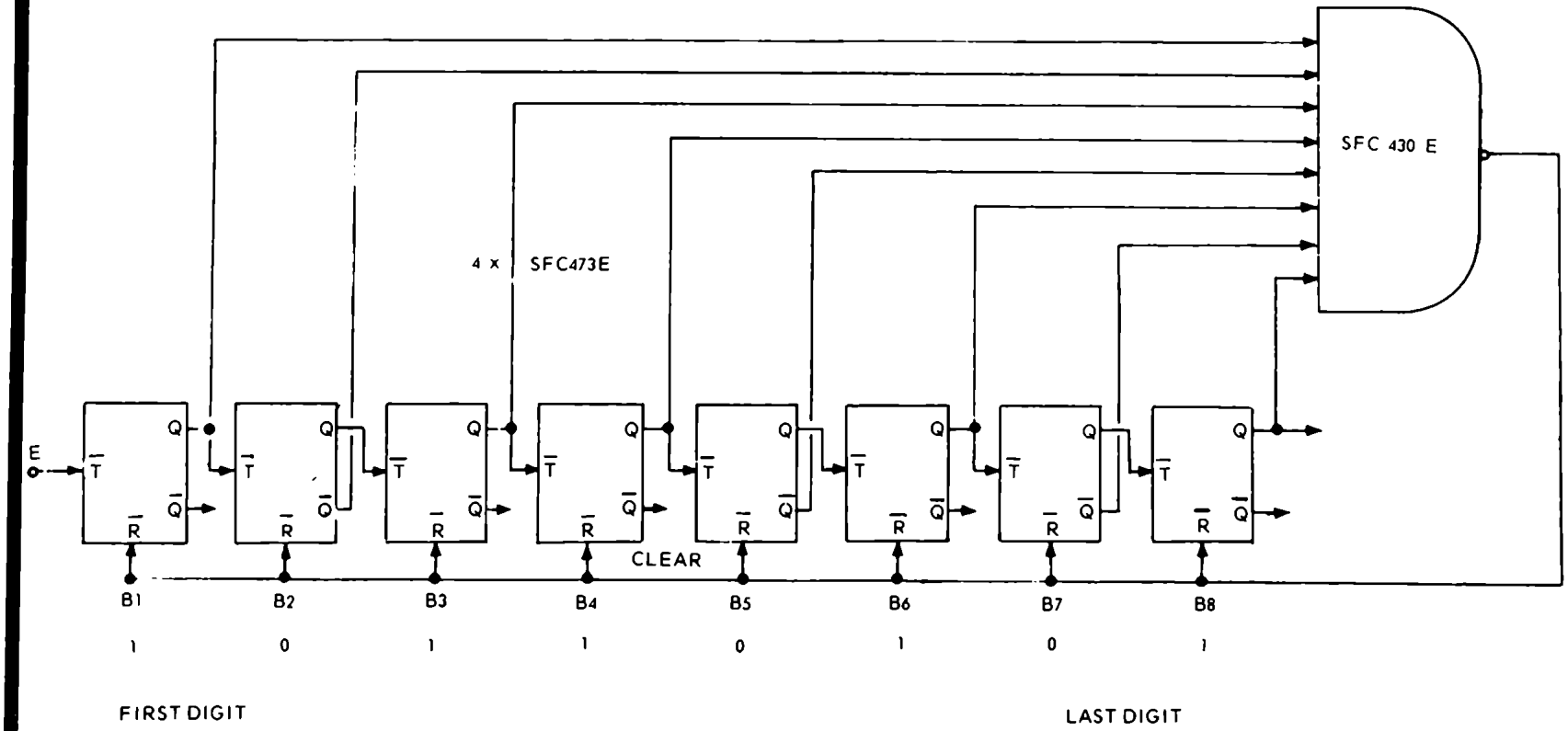


Fig. 5-32. Preset 8-bit binary counter using five integrated circuits.

counting an 8-bit binary number. Thus, the "ordinary" number, 173, corresponds to binary 10101101, meaning 1 on the Q outputs of B1, B3, B4, B6 and B8 and on the $\overline{Q}$ outputs of B2, B5 and B7. These outputs are connected to the eight inputs of a NAND. If the NAND output is tied to all $\overline{R}$s, it will transmit a clearing pulse at the 173th count. This pulse can, of course, initiate any desired action.

To set up a decimal counter (if only for the sake of the easy reading display), at least four scales-of-two are put together and the number of states is reduced from 16 to 10 by suitable back-couplings, as shown previously. However, the manufacturers market complete decimal counting decades composed of four J-Ks and two 2-input NAND gates (SFC490E of SESCOSEM), arranged on a single silicon chip. There also is a decoder (SFC441E) which converts the signals on decade outputs A, B, C and D to digits from 0 to 9 to be applied to the corresponding electrode K0 to K9 of a Nixie-type numerical display tube (Fig. 5-33). This last integrated circuit even contains the 10 high-voltage silicon transistors needed to drive

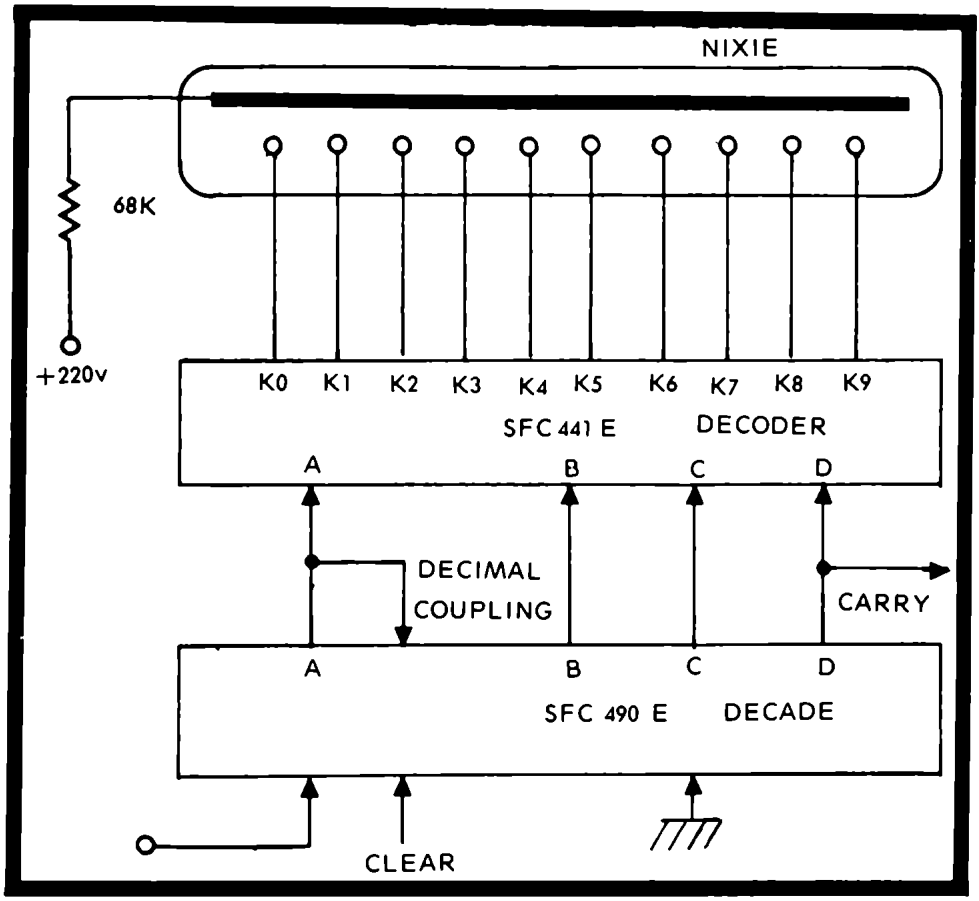
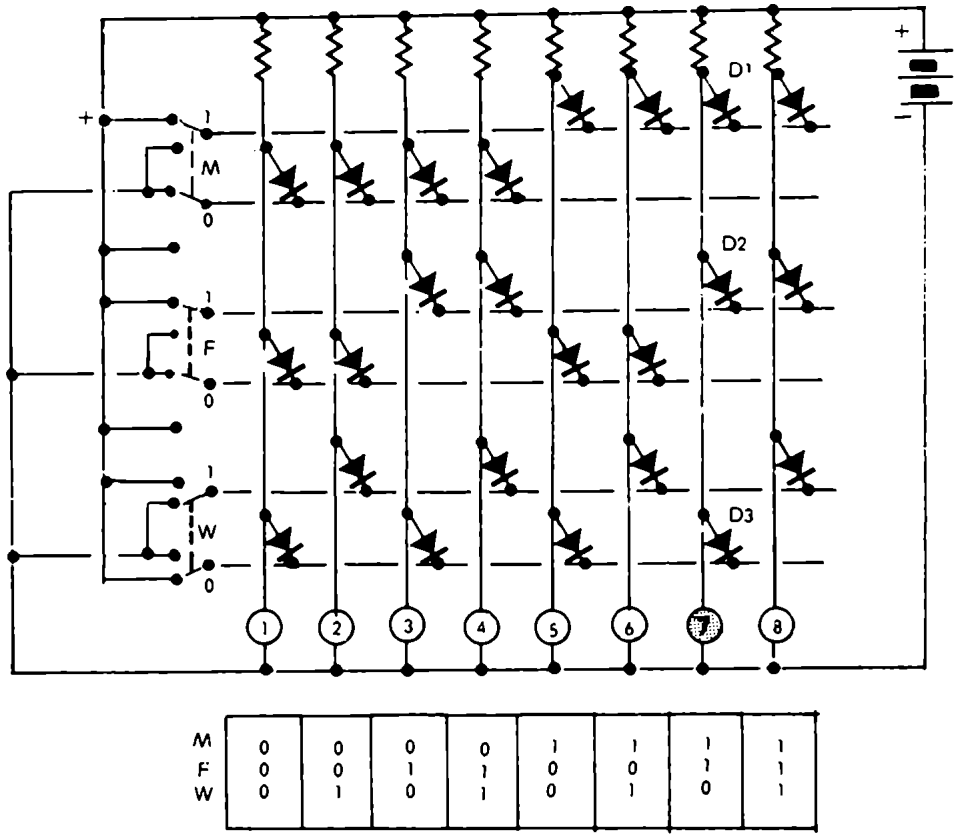


Fig. 5-33. Connection diagram for a single-chip J-K decimal counter with an integrated decoder designed to drive a Nixie-type display.



M	0	0	0	0	1	1	1	1
F	0	0	1	1	0	0	1	1
W	0	1	0	1	0	1	0	1

Fig. 5-34. Diagram of a 3 x 8 diode matrix.

the cathodes. The tenth counting pulse triggers the tens decade in a way similar to the method used to trigger the units decade, and so on. It is not difficult to put together counter and decoder integrated circuits and Nixie tubes to get a custom-made decade counter, and the development of all kinds of digital instruments of course has been mightily spurred by this convenience.

DECODERS FOR DISPLAY DRIVING

The four outputs, A, B, C and D of a decade deliver voltages that vary simultaneously in a manner unrelated to the required electrode voltages of a Nixie tube corresponding to the same count. Therefore, a decoder is needed to convert the counter output into Nixie driving voltages. Such a commercial decoder was mentioned earlier; let us see now how it works.

Previously, we used the example of an automatic drillpress control to illustrate the process of logic diagram tracing. It may be necessary or desirable to identify the situation arising with variables M, F and W occupying given states. The diode matrix in Fig. 5-34 will decode the in-

formation contained by the states of the variables and light one of the eight lamps, 1 to 8 (and only one), to display the situation number.

Such a matrix is composed of vertical bus bars which make available the decoded outputs and horizontal bus bars for applying the inputs. The crossing of the conductors are noncontacting; but unidirectional current flow can be provided by connecting diodes between suitable horizontal and vertical rows. On the diagram, the E-J circuits are figured by DPDT switches and the switch position simulates the circuit state. On the diagram, bus bars M1, F1 and W0 are tied to +; therefore, they will shunt the series resistor and light lamp No. 7, the only one that is not connected by a diode to a negative line. This means that the motor is running, the cutting fluid flows, but there is no work feed, an unacceptable situation.

To simplify this diagram, only one diode is represented to show its polarity, and the crosspoints to be fitted with diodes are circled in Fig. 5-35. Similarly, in a counter decade, a matrix can be used to display the positions of the EJs, the only

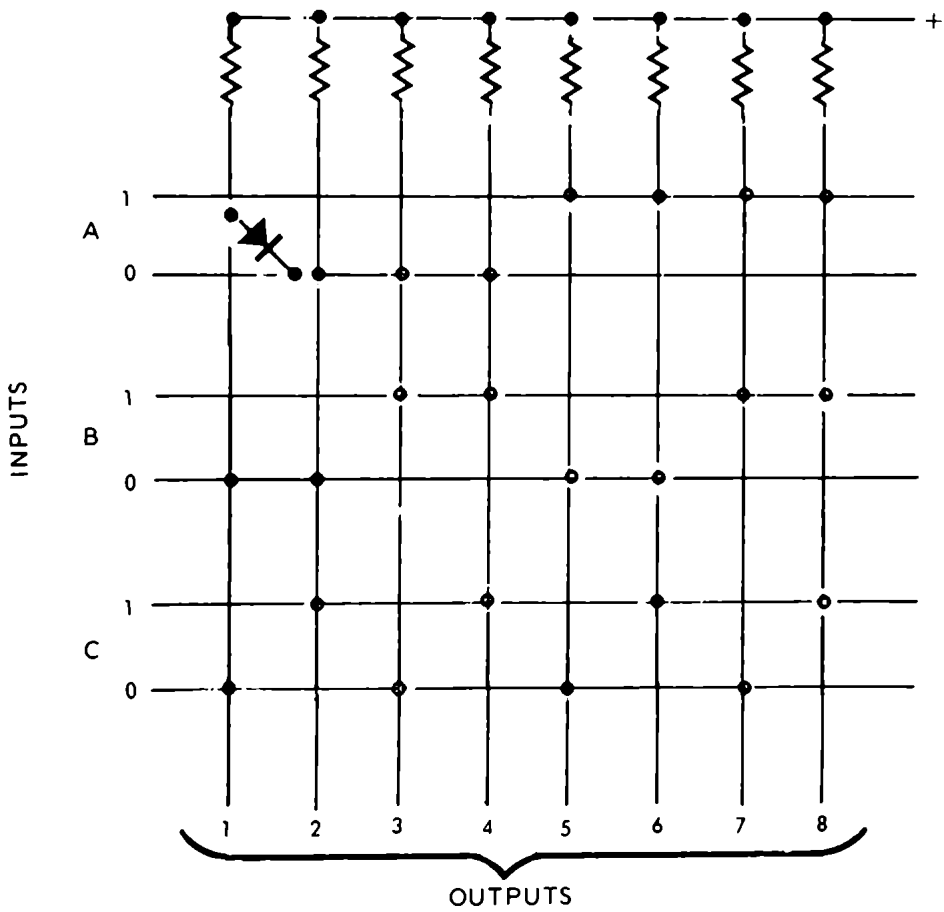


Fig. 5-35. Simplified representation of a diode matrix.

means of reading the count. The operation of the decimal counter in Fig. 5-35 is described by the truth table, Table 5-2. The diagram in Fig. 5-35 is easily modified to successively light the 10 lamps of a numerical display such as shown in Fig. 4-23, featuring convenient in-line reading.

Of course, this matrix principle holds true for Nixie-type numeral display tubes. But since they are voltage-controlled devices, Nixie tubes require special driver circuits as shown in Fig. 5-36, using high-voltage silicon transistors. Using a PNP transistor decade and NPN drivers, decoding has to be accomplished at the 0-state coincidence by means of OR gates (DTL). All transistors having at least one diode input tied to level 1 (-10v) are blocked. Resistor values and base bias (+3v) are chosen for transistor saturation to take place with all OR

	E-J 1		E-J 2		E-J 3		E-J 4	
Situation	$\overline{Q}$	Q	$\overline{Q}$	Q	$\overline{Q}$	Q	$\overline{Q}$	Q
0	0	1	0	1	0	1	0	1
1	1	0	0	1	0	1	0	1
2	0	1	1	0	0	1	0	1
3	1	0	1	0	0	1	0	1
4	0	1	0	1	1	0	0	1
5	1	0	0	1	1	0	0	1
6	0	1	1	0	1	0	0	1
7	1	0	1	0	1	0	0	1
8	0	1	0	1	0	1	1	0
9	1	0	0	1	0	1	1	0

Table 5-2. Truth table showing the operation of the decimal counter in Fig. 5-36.

inputs at level 0. This grounds the corresponding Nixie cathode, illuminating the numeral. With its driver transistor nonconducting, the cathode potential determined by the voltage divider is too high for the tube to strike. (The 1-meg resistor shunting every transistor avoids excessive voltage buildup across the nonconducting semiconductor device.)

The seven bar-type display in Fig. 4-24, of course, needs a decoder, too to allow for a decade readout. The special matrix can be devised, considering Fig. 4-25, for the bar arrangement correlated to every digit.

STATIC CONTROLS USING MAGNETIC DEVICES

There is still another static switching technique using magnetic instead of semiconductor devices; the operating principle is based on switching the magnetic flux in a core

made of a special ferromagnetic alloy and fitted with several independent windings.

The magnetic characteristic (called a hysteresis loop used here) has a rectangular shape (Fig. 5-37); this graph represents the induction B of the material in terms of the applied magnetic field H. If a positive polarity field is applied to a magnetically new core, the operating point passes from 0 to C. (This is the first magnetization cycle; barring thorough demagnetization, Point O will never be reached again.) Because this is a typical high-permeability material, B rises steeply with H. Suddenly, there is a marked saturation effect that bends the characteristic to a practically horizontal level; a further increase of H will have no effect on B. (This sudden saturation is typical for magnetically hard materials; soft steel as used for transformer laminations exhibits a gradual

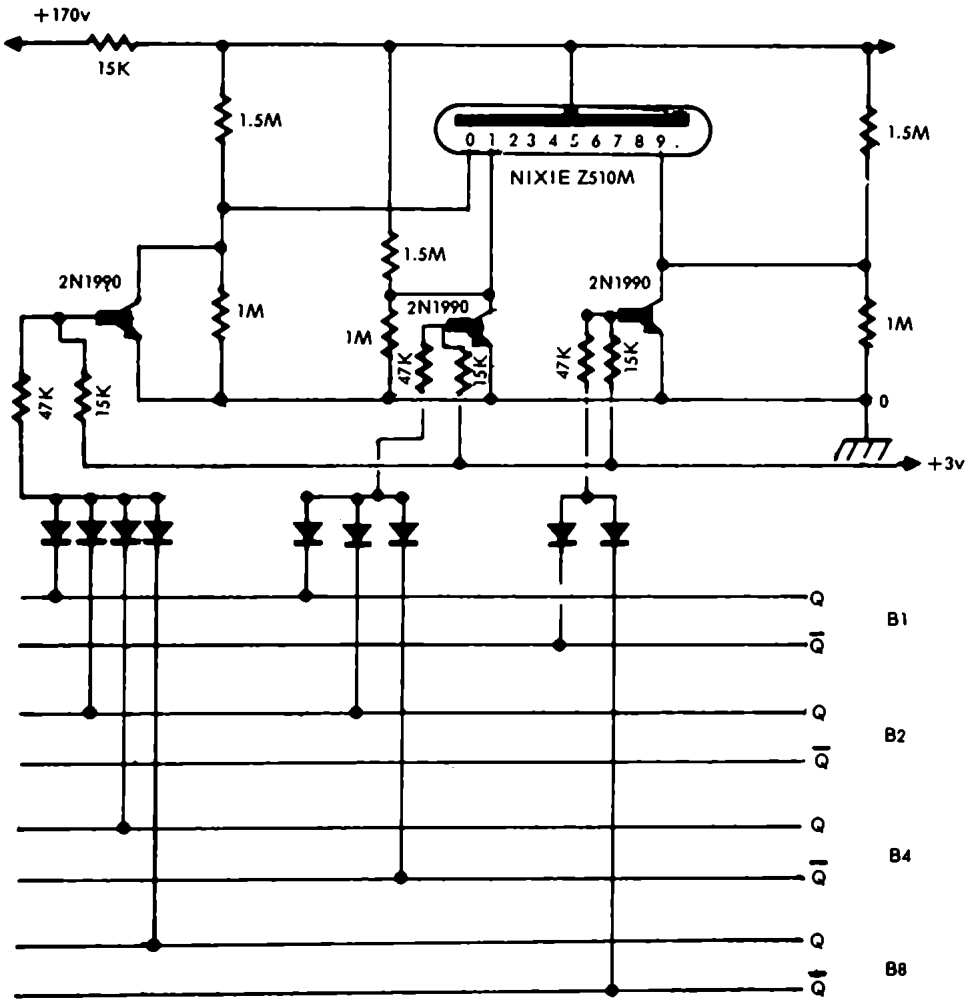


Fig. 5-36. Schematic of a binary decimal coder and silicon transistor driver.

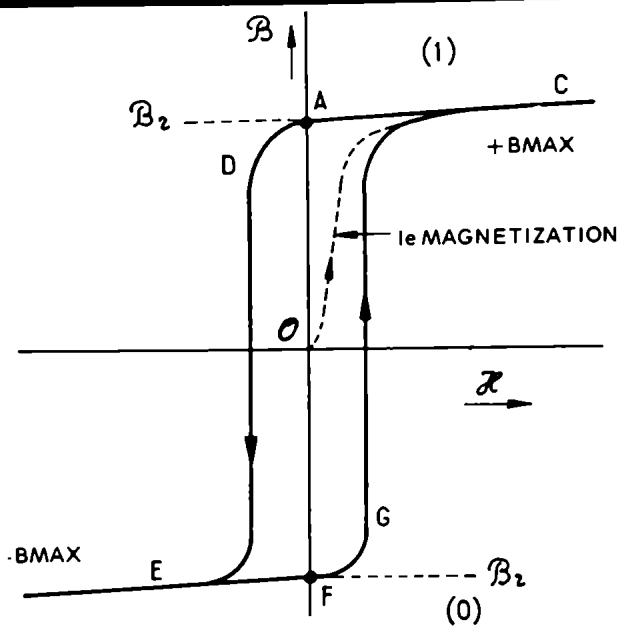


Fig. 5-37. Magnetic hysteresis loop.

bend of the characteristic with no pronounced saturation effect. Such a material is of no use for static switching.)

If the field H vanishes, the operating point shifts from C to A. Because of the pronounced saturation, the no-field residual induction Br practically equals induction Bmax at Point C. Restoring H does not noticeably change matters; however, if a field -H of opposite polarity is set up, the operating point proceeds from A via D to E and settles at F when magnetization ceases, with an induction -Br. While restoring -H, now inoperative, +H will send the working point from F via G and C to A. This magnetic device thus exhibits two stable states (Points A and F) much like a scale-of-two; these states are described by the two polarities +Br and -Br of a same inductance value. So +Br can be identified with 1 and -Br with 0. Since information storage needs no energy here, breaking the power supply does not destroy the stored information.

A device composed of a magnetic core made of a square hysteresis loop material, with three windings A, B and C (Fig. 5-38), can be used to carry out logic operations AND, OR and memory. The signals setting up field H are applied to A and B, and C is the readout winding. A signal that does not reverse Br does not induce a voltage in C. If a signal reverses Br, the sudden magnetic flux reversal induces a voltage pulse across C. To know the polarity of Br, and hence the bit contained, it is necessary only to apply to A or B a signal of known polarity

and to look for a pulse across C to know if the core state was 0 or 1. (This is **destructive readout**; there are nondestructive readout devices retaining the stored information and, thereby, allowing for repetitive interrogation.)

This obviously is the memory function. An OR gate is obtained when only one signal, A OR B, is necessary for flux reversal. If the amplitude of a single signal is inadequate for switching, an AND gate is obtained, for signals A AND B must be present simultaneously for the added fields to operate the tip over. It is also possible to use signals A and B of the same amplitude but opposite polarity so that there is no field if A AND B are present simultaneously.

Since a winding can boil down to a single conductor traversing the core, compact and economic mass memories can be built using many tiny toroid cores (⅛- to 1/16-inch outer diameter) woven in a **ferrite core matrix**, according to Fig. 5-39, to form a memory plane. There are horizontal and vertical input wires A1, A2, A3 and B1. B2, respectively, and a common readout wire, C, positioned diagonally. Information is entered by the coincidence of two signals applied to one horizontal and one vertical wire (A2 and B1, for instance). Both signals must be of sufficient amplitude (AND function) to cause a flux reversal in the single core traversed by lines A2 AND B1.

Since the readout is destructive here, the destroyed information can be restored by reading-in a delayed pulse each time a pulse has been read out. This is done by running special readin wires in parallel with input wires A and B; thus, every core now is traversed by five wires.

A static control, and even a digital computer, can be built using only saturable ferromagnetic cores, with gain provided

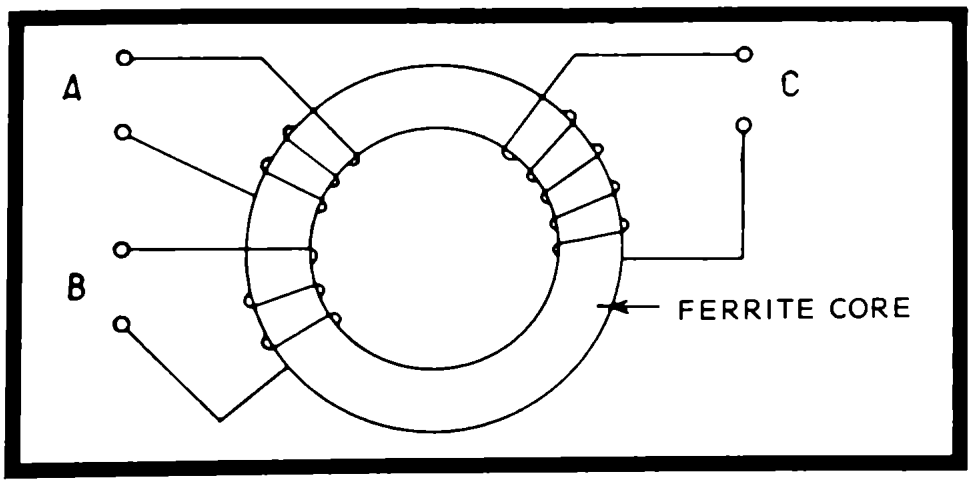


Fig. 5-38. Schematic of a ferrite switching core with three windings.

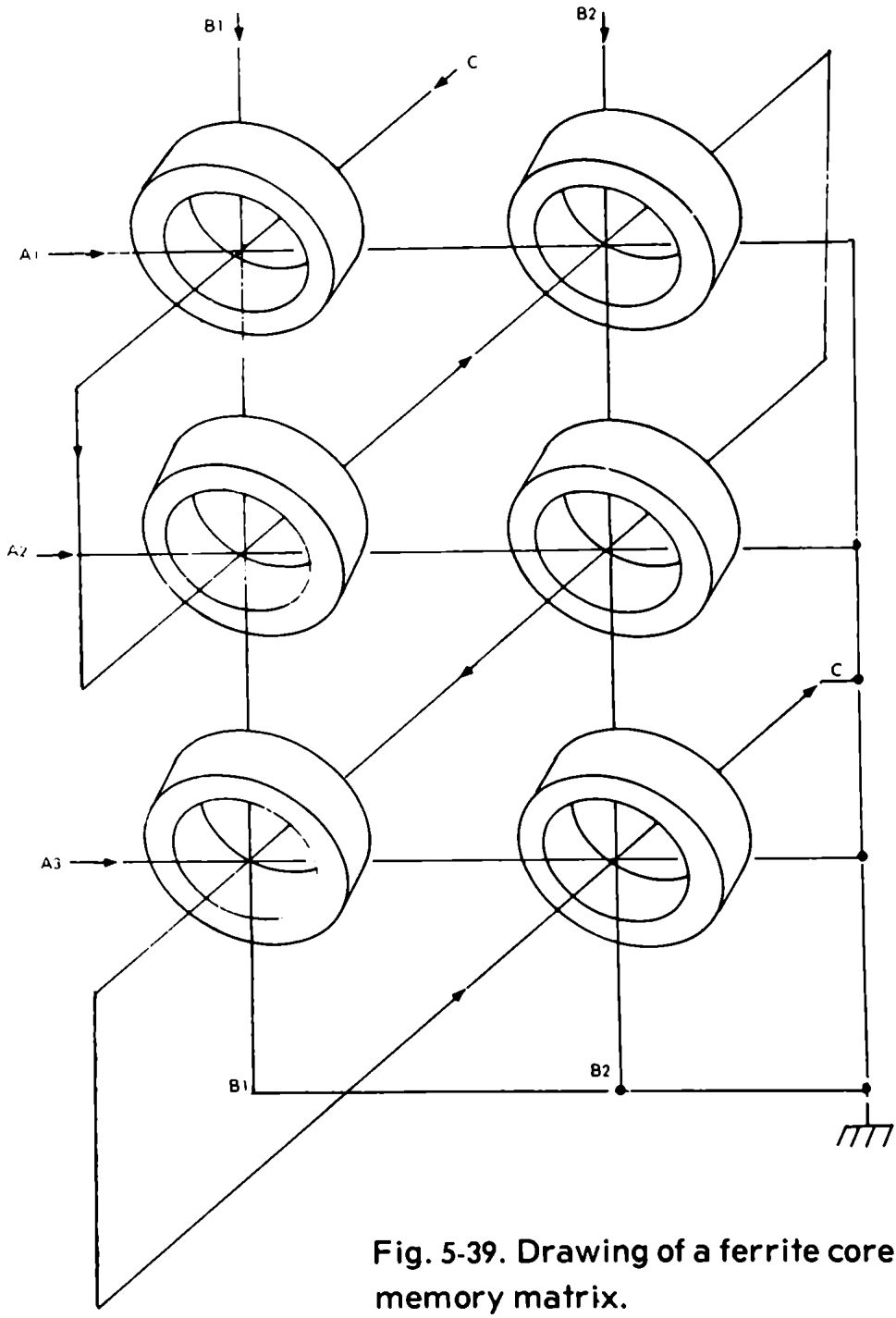


Fig. 5-39. Drawing of a ferrite core memory matrix.

by the Ramey principle of flux resetting used in magnetic amplifiers. The cores can also perform the logic functions, while associated transistors provide for the required gain. In the circuit in Fig. 5-40, the base of a PNP transistor is biased positively through readout winding C, which blocks the transistor. An induced negative pulse will result in a positive output pulse, serviceable for driving other cores.

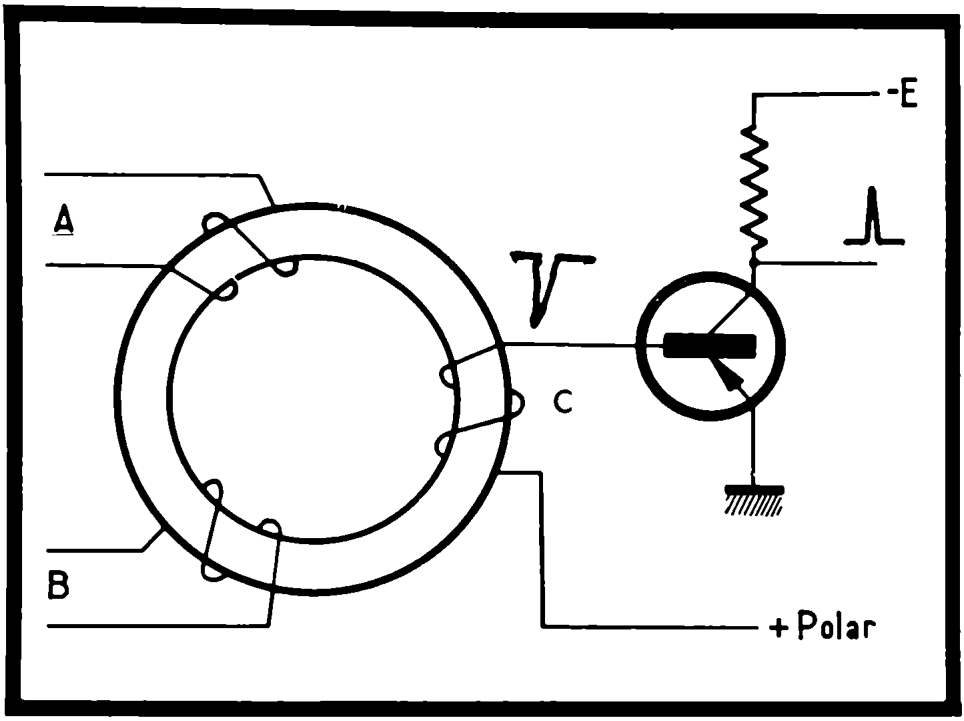


Fig. 5-40. Schematic of a magnetic core signal amplifier.

A transistor can also be used to generate feedback to accelerate the tip over. With windings C and D connected to the base and collector of a transistor, the circuit in Fig. 5-41 looks like an oscillator; the transistor, however, is blocked by a positive bias, and the (positive) feedback is insufficient for sustained oscillations. On arrival of a pulse, the transistor

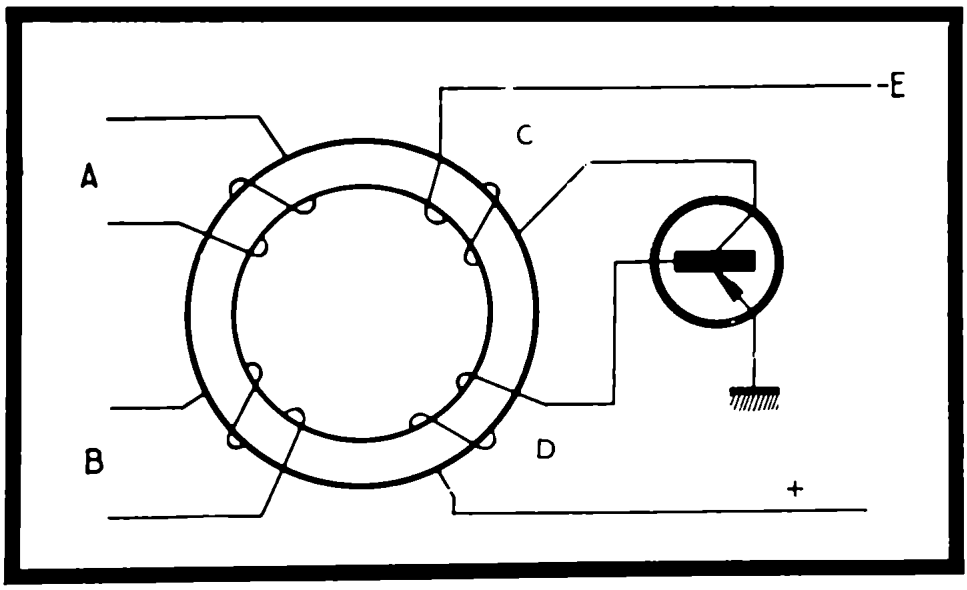


Fig. 5-41. Magnetic amplifier circuit with transistor feedback.

conducts, and the differential core permeability dB over dH, which is low at saturation, takes a very high value, resulting in a powerful feedback until the switching process is completed. This allows us to work with very short pulses at repetition frequencies of up to 100 kHz.

Several manufacturers offer static controls based on magnetic amplifier principles, such as Westinghouse's CYPAK system composed of matchbox-size potted units that plug into wired panels just like semiconductor controls. Recent semiconductor circuits, however, allow for much higher operating speeds and offer utmost flexibility; also, integrated circuit prices have come down to a point where competition is difficult for magnetic controls. So it seems likely that semiconductor logic will gradually supersede magnetic circuits in all but power applications (Chapter 8).

ANALOG-TO-DIGITAL CONVERSION

In the foregoing, we have dealt again and again with the two fundamental kinds of electronics: continuously varying, or **analog** signals, and on-off, or **digital** data. If a continuously variable process parameter is to be sensed and handled by digital methods and equipment, an analog-to-digital converter (ADC) must change the expression of the information from analog to digital.

The angular-position digital encoder is such an ADC. Its key part is an encoder disc attached to the shaft whose angular position is to be monitored. This disc is made of a transparent material such as glass, partly recovered by a concentric arrangement of opaque sectors. (High accuracy encoder discs are made by photographic reduction of large-scale artwork.) One side of this disc is illuminated by a light source, and on the opposite side is arranged a radial row of photocells in a way to illuminate the cells unless the light beam is intercepted by an opaque sector (Fig. 5-42). An illuminated cell equals 1, and 0 stands for a dark cell (or vice versa).

Fig. 5-43 is a detailed view of such an encoder disc. Four rings of transparent and opaque sectors read by four photocells deliver 4 bits, dividing each turn into 16 angular positions. This generally is too coarse a division; with a fair number of sector rings and photocells, a quite high angular definition is obtained: with **n** rings and cells, the circumference is divided in 2n parts. Matchhead-sized diode-connected phototransistors are used for readout. Plastic light pipes are convenient for connecting the phototransistors to the encoder disc. Instead of the photoelectric device, fine contact brushes wiping conductive sectors separated by insulating

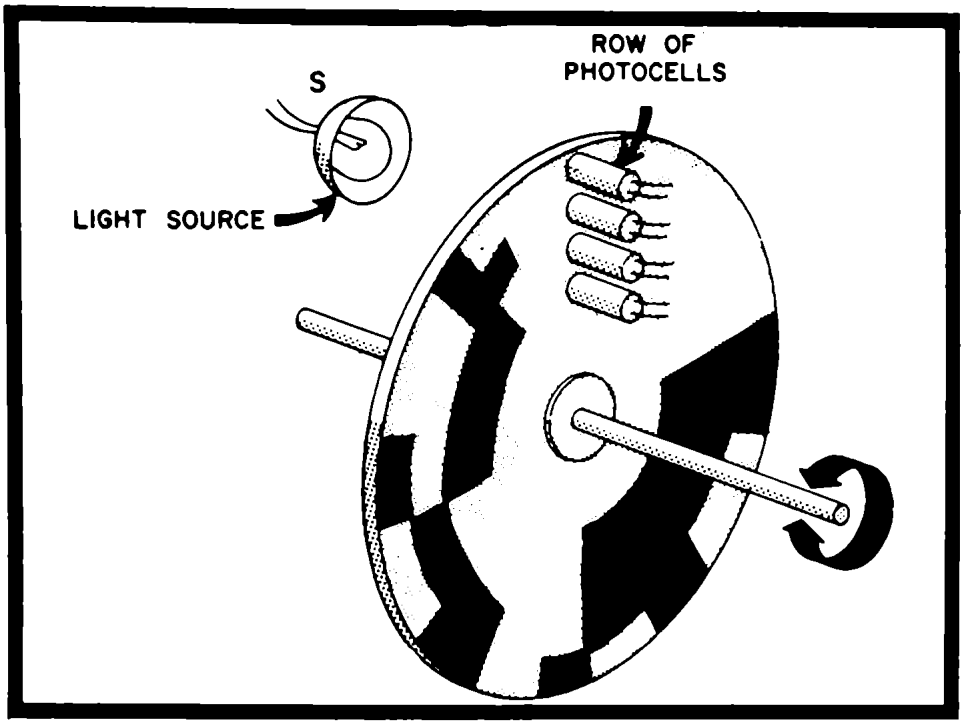


Fig. 5-42. Photoelectric analog-digital angular position encoder.

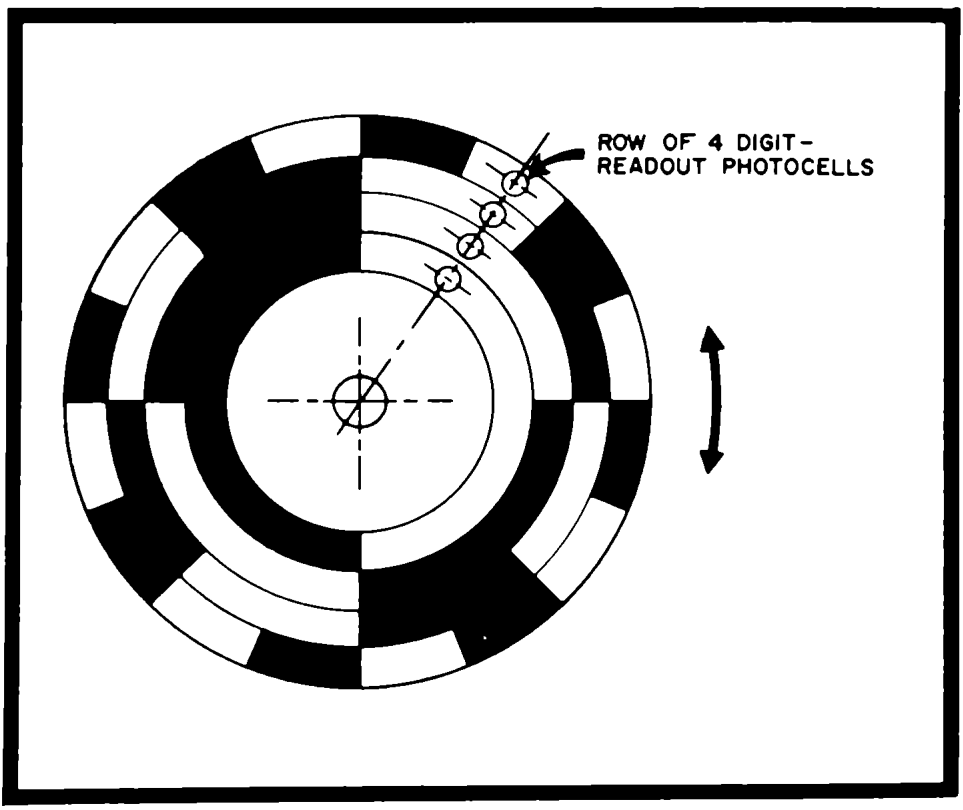


Fig. 5-43. Drawing of a 16-position angular-digital encoder disc.

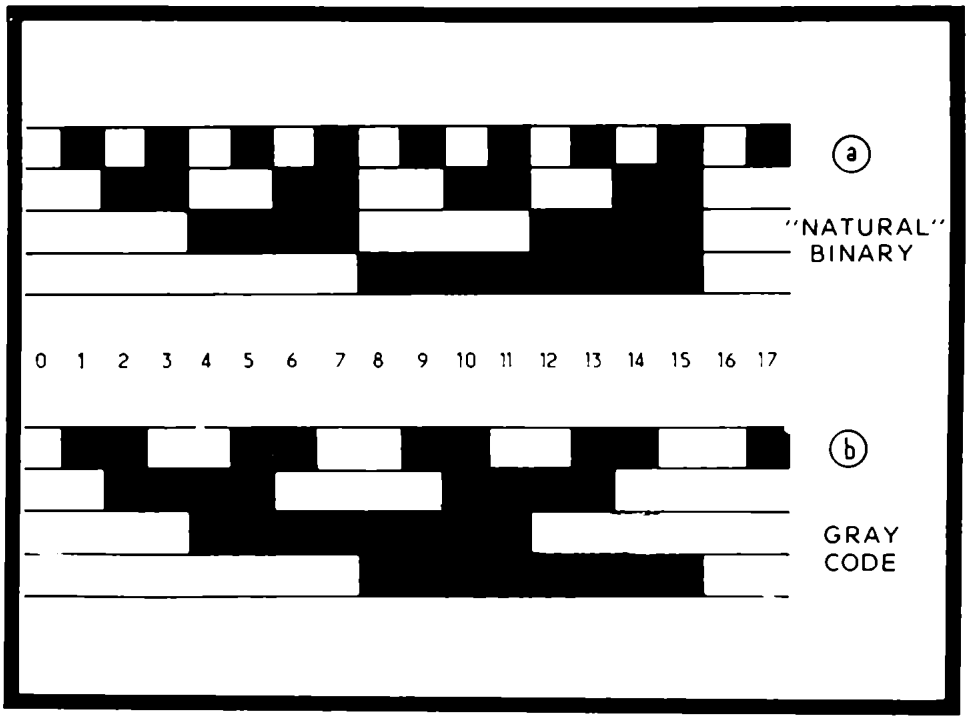


Fig. 5-44. For analog-to-digital conversion the Gray Code (B) is preferred to the natural binary code (A).

areas are serviceable as well. Such discs are fabricated by printed-circuit techniques. The circular divisions can also be arranged longitudinally for encoding linear displacements of machine tool carriages and similar applications.

The correlation between the angular position (or linear displacement) and the binary readout is defined by a **code**. The existence of several codes denotes that for a given application, one code is better suited than another. The encoder disc in Fig. 5-43 is based on the so-called "natural binary" code shown linearly (stretched out) in Fig. 5-44A. Notice that between 15 and 16, all digits change simultaneously, and even between 7 and 8, three out of four digits change. This simultaneous variation can be troublesome, for when shifting from 15 to 16, all photocells may not be masked or all contacts may not be made at the instant the readout is made; therefore, an erraneous digit combination may result.

This risk is reduced by varying a single digit at a time, using a so-called **cyclic code** such as the **Gray code** (Fig. 5-44B). Comparing coding strips A and B, where the black segments stand for 1 and the white spaces for 0, you will notice that for the Gray Code, there nearly always is an overlap of the 1 states between two decimal numbers, thereby minimizing error possibility. Table 5-3 shows the differences between these codes.

Decimal	Natural Binary Code	Gray Code
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100
9	1001	1101
10	1010	1111
11	1011	1110
12	1100	1010
13	1101	1011
14	1110	1001
15	1111	1000
16 (5 0)	0000	0000

Table 5-3. Comparison of the natural binary and Gray Codes.

PROGRAM CONTROL

It takes a lot of skilled operator work hours to set up a machine tool to machine a given workpiece. But such a machine rarely turns out a single item permanently; so the skilled mechanic must again and again reproduce a former setup, and if the new production series is a short one, this is costly and time-consuming.

Therefore, it is fortunate that a particular setup, once completed, can be **memorized** on a suitable information-retaining support, so that a later setup can be started and completed automatically by introducing the once recorded setup information. To pass on to another workpiece run, it then is only necessary to introduce another "program."

Industrial processes can be described in terms of sequences, one thing being done after another; such a sequence is called a **program**. As machines are controlled by binary decisions, the program contains the digital information such as holes punched in cards or paper tape standing for 1, and no holes for 0. A specific position— the **address**—is assigned to every bit of information, for the reading device has to know where to find it. Punched cards have an edge when the company has business computers on hand that can be used for

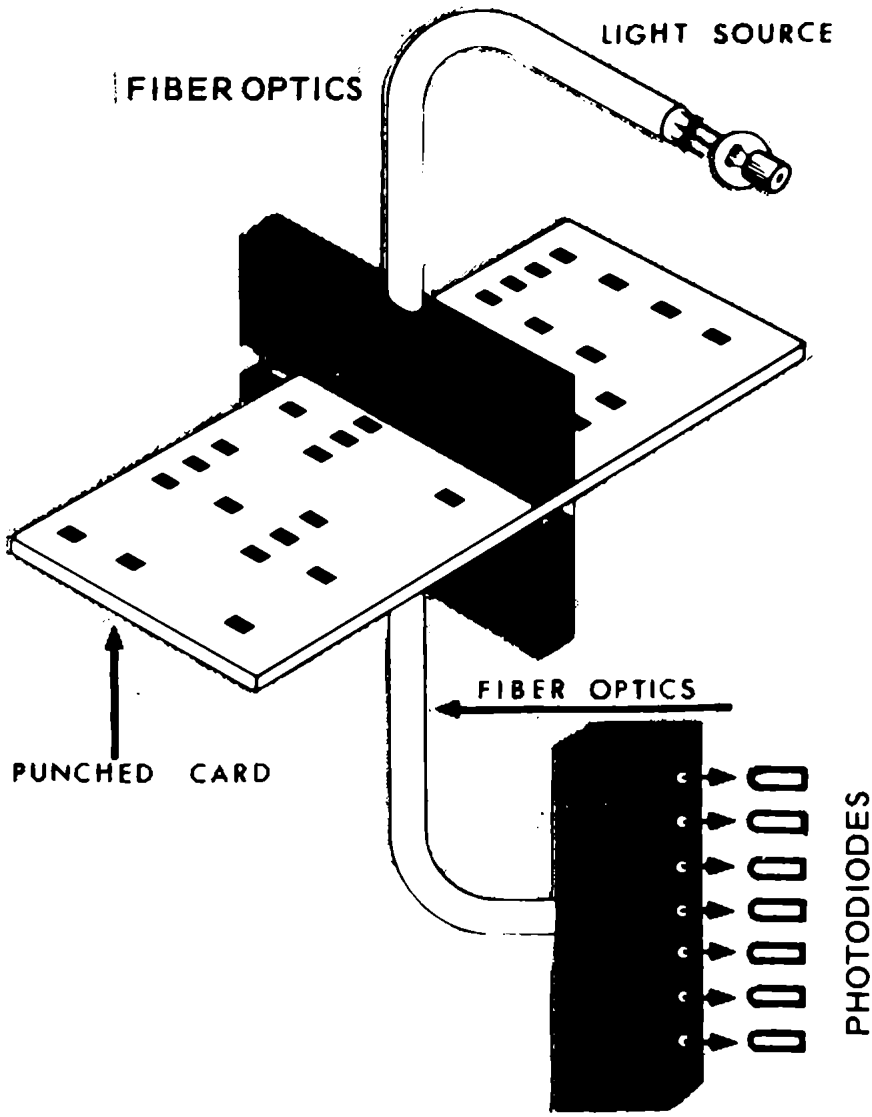


Fig. 5-45. In this drawing, a punched-card reader uses optic fibers as light ducts.

punching, processing and checking the programs. Punched tape is somewhat faster than punched cards and is easier to store. Magnetic tape (up to 2 inches wide) is used in special multiple-head industrial-type recorders dealing only with pulses.

In punched tape, information is contained in a number of columns read by an equal number of contacting feelers or phototransistors; optic fibers are frequently used for hole position illumination and to connect these positions from the opposite side of the tape to the phototransistor (Fig. 5-45).

Four-column punched tape is frequently used, with a median column of smaller, regularly-spaced holes for the tape transport (Fig. 5-46). The holes are further positioned in regularly spaced rows, each row containing the instructions for one process step; for instance:

Rotate drillstock no. 3 into working position;
Set tool feed at 1.5 inch per min;
Set drilling depth to .25 inch;
Position workpiece by preset dogs 2D and 5G.

The tape comes to rest at every row to enable the reading devices to extract the instructions contained and to carry out the various orders. A reporting-back device is associated with every positioning device or control, and it is only after accomplishing all the operations prescribed that the tape can proceed to the next row, where the readers may find the following instructions:

Energize the drillpress motor;
Energize the drillstock feed;
Stop the motors on reaching the prescribed drilling depth;
Raise the drillstock to its rest position after completing drilling.

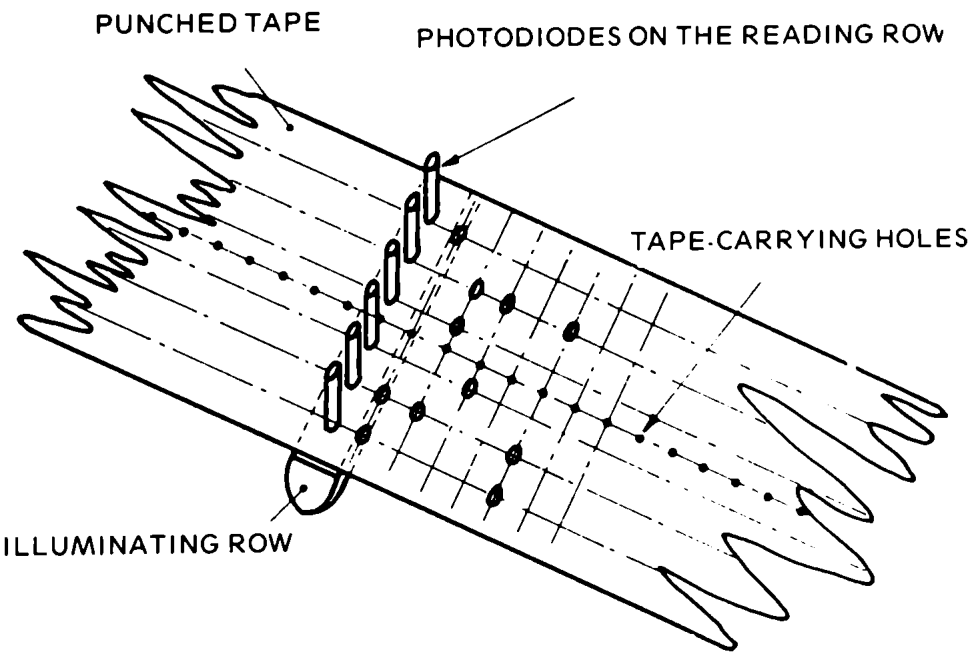


Fig. 5-46. Drawing illustrating the operating principle of a punched-tape reader.

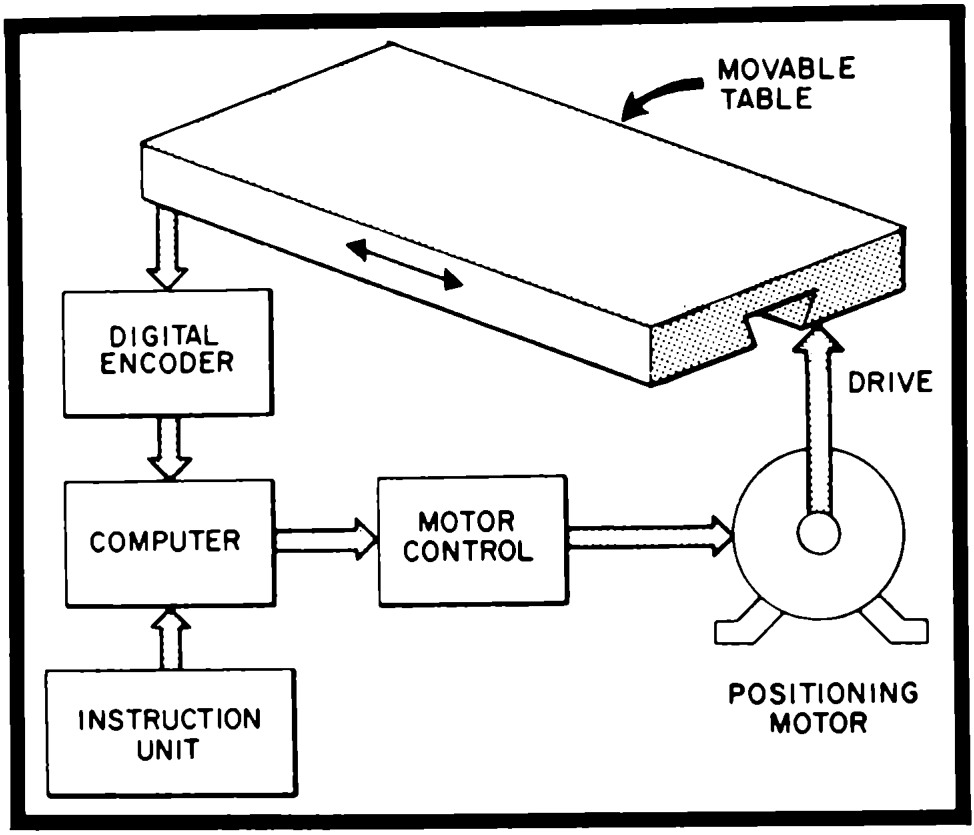


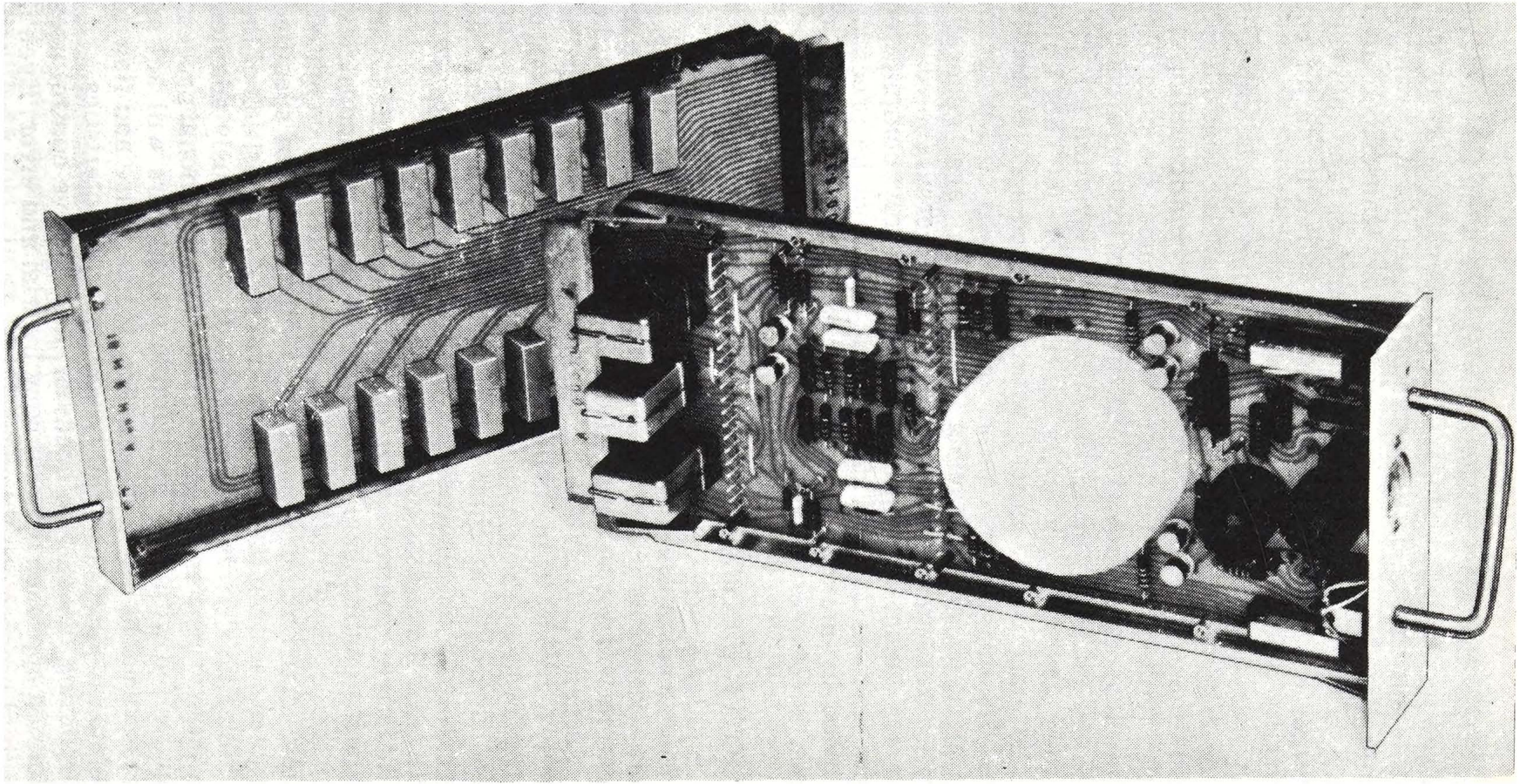
Fig. 5-47. Operating diagram of a numerical positioning system.

NUMERICAL POSITION CONTROLS FOR MACHINE TOOLS

There are two basic types of digital controls. In the coordinate positioning system used in jig borers and vertical turret lathes, the displacement of a movable table along two coordinate directions XY is operated by feed screws driven by servo systems according to programmed instructions. This allows for mass production of precision-drilled workpieces, for instance.

The operation of such a numerical position control is diagrammed in Fig. 5-47. A digital control computer gets its instructions from an input unit fed with punched cards or punched or magnetic tape. On the other hand, the actual position of the table is transmitted to the computer by an analog-digital converter. Comparing the actual and desired positions of the table, the computer issues suitable signals to the motor control to drive the table to the required coordinate point. (The system shown is for only one coordinate.)

As in all closed-loop systems, the true position is continuously reported by the ADC to the computer, and the motor



Analog (left) and digital drawer-type plug-in modules.

stops when the ADC information agrees with the position required by the instructions. This approach does away with costly templates and time-consuming tool setups. A drawing is used to program the position information, transcribed after coding on a medium such as punched cards or tape (punched or magnetic), ready to use and easily stored in an ad hoc library for a further production run.

A second type of numerical control, more sophisticated, is used in contouring systems where the computer continuously controls the path of the cutter or tool. In automatic milling of a circular groove in a workpiece, for instance, the computer continuously calculates the coordinates of points around a circle or radius as well as the maximum deviation allowed, and then controls accordingly the table feeds for both directions to machine the groove within the required tolerance limits. The instructions containing the step-by-step program and incremental positioning information are generally recorded on magnetic tape. Such a truly universal control will allow for automatic milling of geometrically complicated shapes, such as turbine blades, that could not be precision machined by manual control. The equipment required, of course, is complex and expensive.

ADAPTIVE CONTROLS

While a machine stubbornly follows only its instructions, a human operator is (or can be) prompted by a search of perfection: he will not be satisfied with accomplishing a task, he tries (or can try) to do it otherwise to get a better result; he actually learns by trying to perfect the operating conditions of the system he controls. Much research is now going into so-called "learning machines" that not only carry out an assigned task, but unceasingly try to optimize the process. Such a control adapts itself to a changing situation.

Adaptive controls are particularly attractive for monitoring continuous processes, in the chemical industries, for instance. Raw materials and ingredients are weighted, treated, mixed and processed to yield a product of specified characteristics, but because of the varying composition of the raw materials delivered, the process is continuously adapted to maintain consistent product quality. A conventional process computer fed with data by suitable transducers will act on process variables such as pressure, temperature and flow to maintain constant specified processing characteristics.

Opposing this slavish behavior, an adaptive control uses the human approach of optimizing by trial and error. Even if the actual process conditions result in an acceptable product

in terms of quality and cost, the computer does not know if still better conditions could be set up, so it deliberately introduces a slight change just to see what it does. Let the process be represented by a diagram such as Fig. 5-48, where the highest yield corresponds to a pressure P (Point A), and let B be the actual operating point, corresponding to a slightly lower yield. Since the control does not know if a higher yield is obtained by increasing or by decreasing the process variable, it may try a pressure reduction, leading to Point C. After an unavoidable time lag, the change appears unsatisfactory because of a lower yield. So the control tries the opposite way and increases pressure to Point A. Since the yield is higher, a further increase to Point D is tried. But this, again, decreases the yield, so the next action is a pressure decrease towards B. Thus, the control is always busy trying for better conditions, and as the various parameters of such a process line are varying and

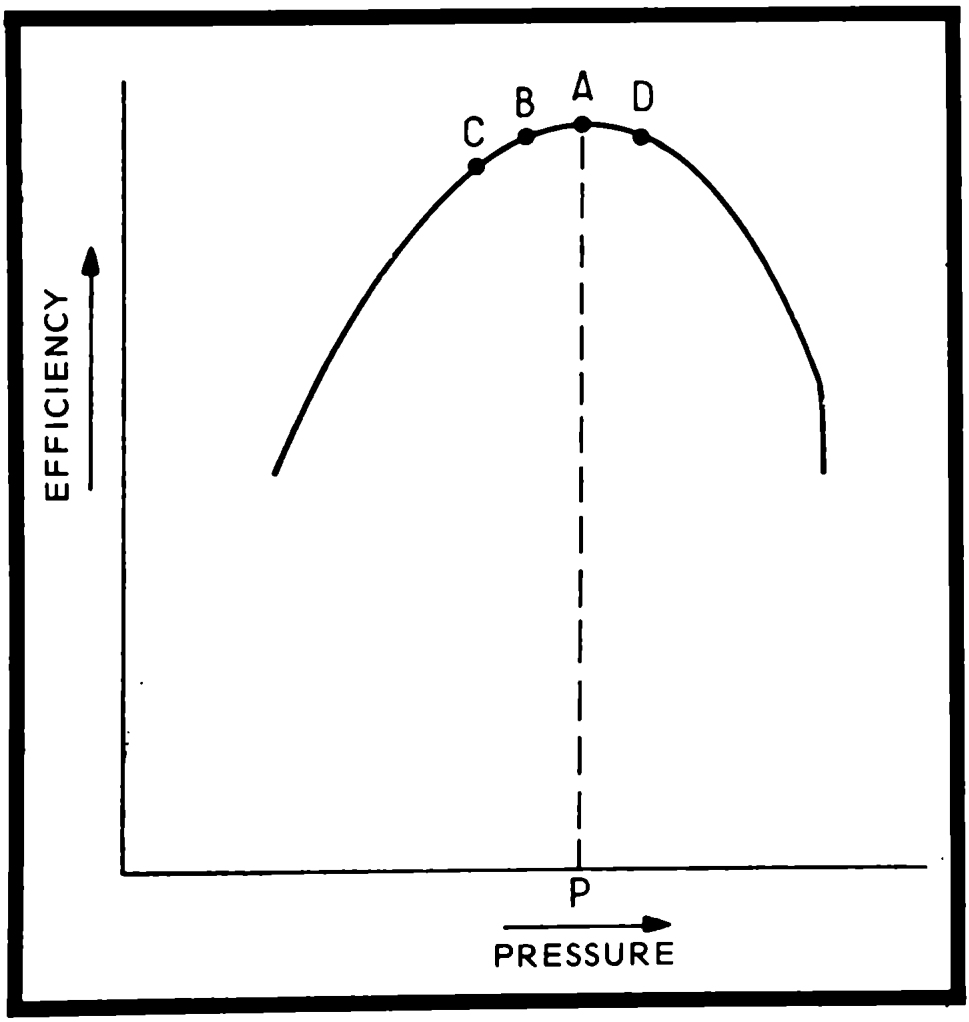


Fig. 5-48. Adaptive controls are always trying for better results as this curve illustrates.

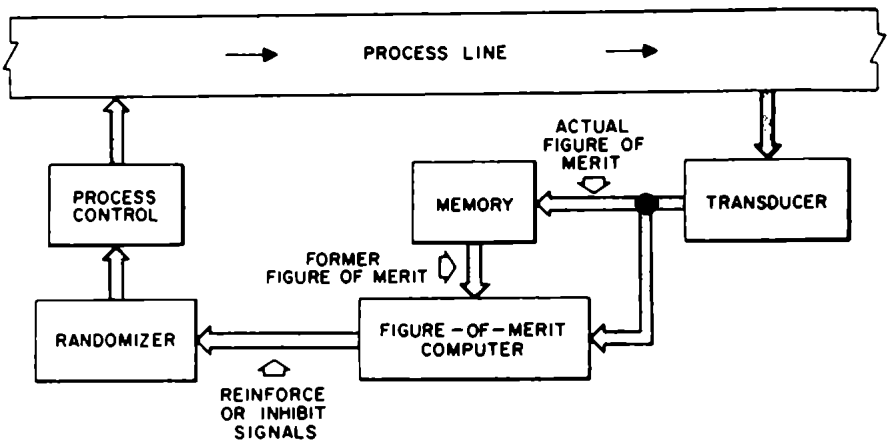


Fig. 5-49. Operating diagram of an adaptive control.

depend on the fluctuating composition of the raw materials, it will never come to rest.

Such an adaptive control system can be illustrated by the diagram in Fig. 5-49. The process line is monitored by a conventional process computer, and the characteristics of interest in the end product are sensed by suitable transducers. An optimum operating condition or figure of merit is defined, and the computer signals are applied to a figure-of-merit computer which calculates the actual figure of merit and compares it with the preceding one stored in a memory. Depending on whether the change proves to be a good or a bad bet, the computer issues reinforcing or inhibiting signals to a "randomizer," resulting in small, but definite changes in processing. The loop is closed by a product reflecting these changes.

In spite of using humanlike trial-and-error methods, the "initiative" of such a system remains limited to the operations assigned and actual "learning" is restricted to the parameters to be monitored. While such a system is complex, it is still a far cry from duplicating the tremendous adaptability of a thinking being.

ANALOG COMPUTERS

Most sensing transducers are analog devices, for they change the measured physical magnitude in a proportionate electric magnitude, its analog. The controls directly using such analog information are analog, too. Digital systems require an analog-digital conversion prior to processing, and

their digital output generally has to be reconverted to analog signals for motor and solenoid control.

Analog computing is a mathematical tool which allows a description of the behavior of any physical system by the study of an electric circuit whose operating equations are analog to those of the physical system. In other words, the behavior of a mechanical system is **simulated** electrically. The mathematical calculus associated with the suspension of a car, for instance, is time consuming and has to be restarted every time a parameter is changed. You can, however, simulate the mass of the car, the force of the springs and the damper action by electric functions, transform the equations in a computer program and directly read the result or display an analog diagram on a scope screen. To modify a parameter value, it is necessary only to readjust a potentiometer. The time gained by this process is easily imagined.

In an analog computer, a problem magnitude x is correlated to a voltage X; the sign of x is reproduced by the polarity of X in DC computers and by an in-phase or out-of-phase condition in AC computers. The maximum voltage is plus or minus 50v or plus or minus 100v in DC computers.

To reproduce X at a suitable scale regarding x, a **coefficient potentiometer** (Fig. 5-50) is used. This reduces applied voltage E in the ratio K equals R1 over R (the very high load resistance can be neglected), and we get output voltage U equals KE. The potentiometers used have resistance values of from 5 to 100K and are fitted with scales calibrated for directly reading K. With helicoidal pots, an accuracy of as high as 0.01 percent is obtained; other models feature an accuracy of 0.1 to 1 percent. The calibrating accuracy, of course, determines the computing precision.

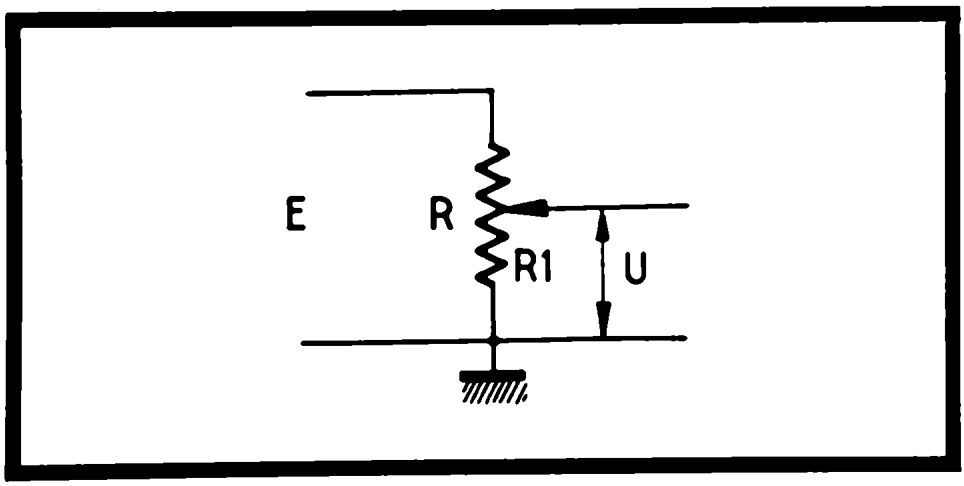


Fig. 5-50. Schematic of a coefficient potentiometer used in an analog computer.

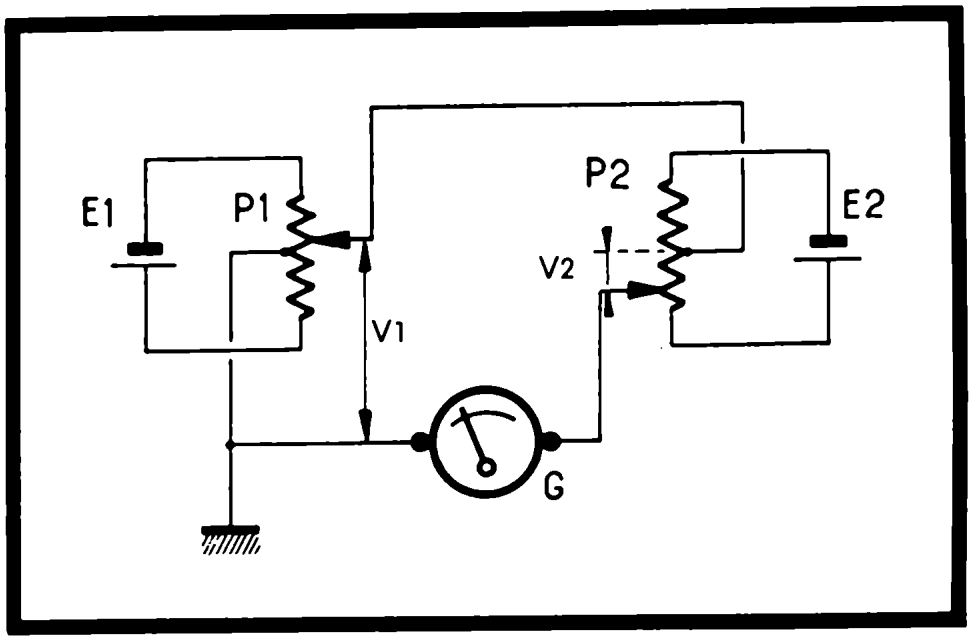


Fig. 5-51. Schematic of an analog voltage adder using two potentiometers.

Analog addition of two voltages, V1 and V2, by means of two potentiometers, P1 and P2, is illustrated by Fig. 5-51, and the result is read on voltmeter G. Due to the use of center-tapped potentiometers, positive as well as negative-polarity voltages can be added.

The center-zero galvanometer in Fig. 5-51 is used in simplified analog computers only because of its limited sensitivity. Transistorized or integrated **operational amplifiers** are standard equipment in analog computers now. The well-known integrated circuit uA702A in Fig. 3-32 (Chapter 3) has the essential features of an operational amplifier: high gain, high input impedance and inverting as well as noninverting inputs. On the other hand, with no input signal, the output terminal has to be at zero potential and stay there; this means neither offset voltage nor drift are allowed.

Terminal 2 of such an operational amplifier, schematically represented in Fig. 5-52, is not only used as a polarity-reversing input, but also to close an inverse-feedback loop via output Terminal 7. An **analog phase inverter** with unity gain is completed by series connecting an impedance, Z1, with Input 2 and closing the feedback loop by another impedance, Z2; unity gain results if Z1 equals Z2. To reproduce a signal again with unity gain, but with its original polarity or phase, noninverting Input 3 is chosen for signal injection, and full inverse feedback is obtained, tieing together Terminals 2 and 7 (Fig. 5-53). This is a so-called "follower" circuit serviceable as an **impedance converter.**

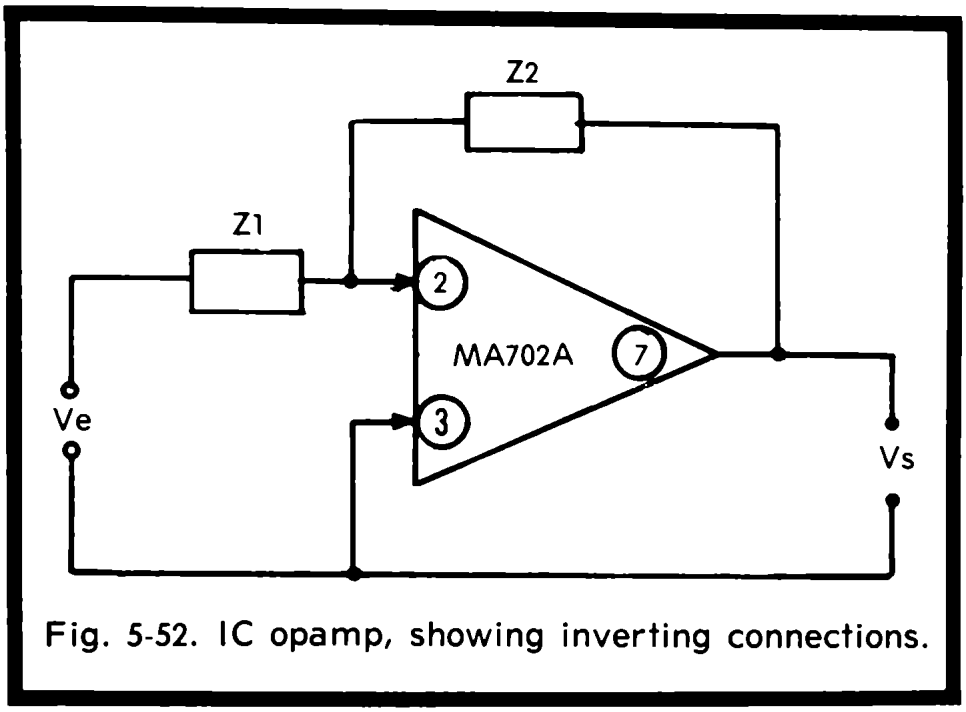


Fig. 5-52. IC opamp, showing inverting connections.

Three voltages, V1, V2 and V3, can be added by injecting them through impedances Z1, Z2 and Z3 into the inverting input (2) of an amplifier provided with an inverse feedback impedance Z (Fig. 5-54). At junction Point M of the four impedances, the sum of incoming currents equals the sum of outflowing currents, that means:

$$\frac{V_1}{Z_1} + \frac{V_2}{Z_2} + \frac{V_3}{Z_3} = -\frac{V_o}{Z}$$

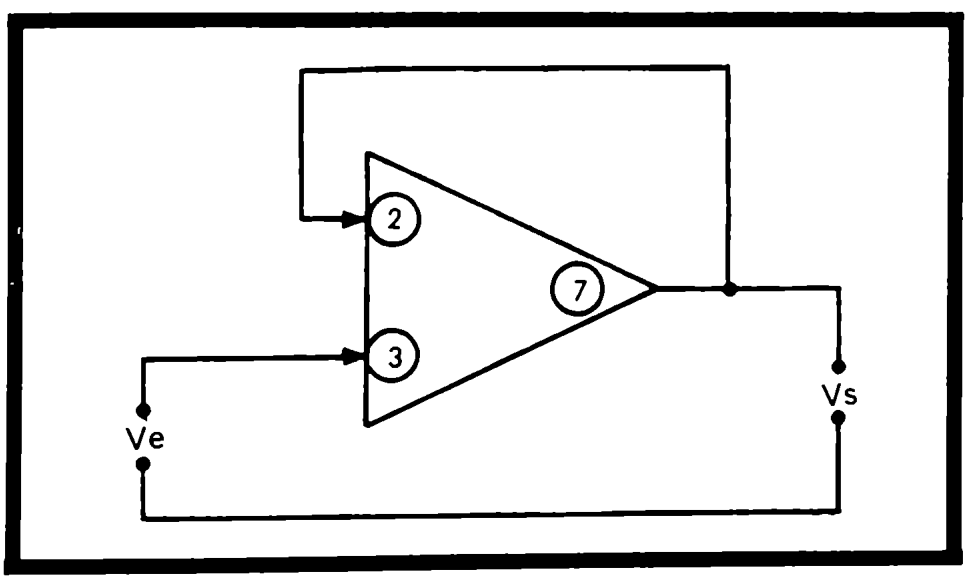


Fig. 5-53. Operational amplifier showing impedance-converter or follower connections.

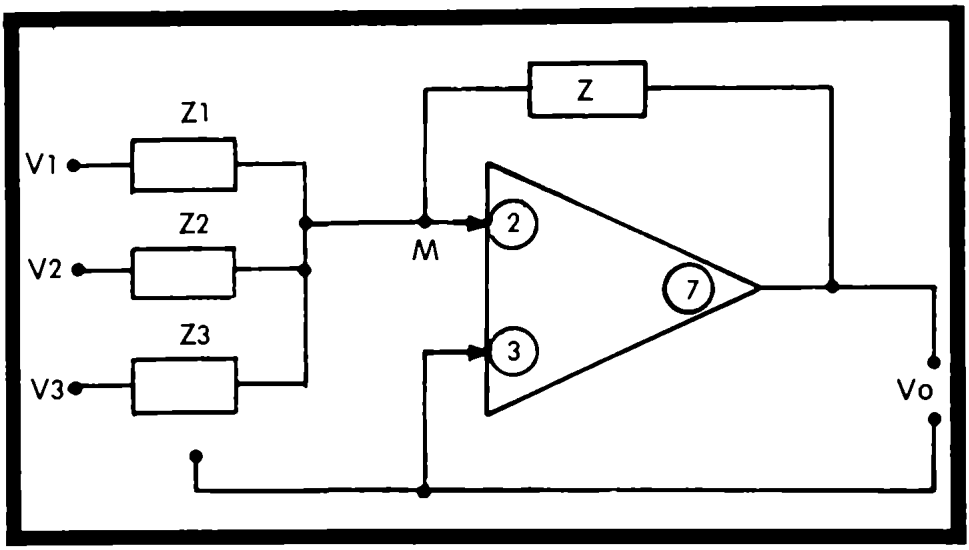


Fig. 5-54. Operational amplifier connected for analog voltage addition.

With Z1 equal to Z2 equal to Z3 equal to Z, we get V1 + V2 + V3 equal to -Vo (the - sign reveals the polarity reversal).

The automatic **adder servomechanism** in Fig. 5-55 is used to solve industrial processing problems; it is based on a similar principle. For error voltage U to be zero, V1 + V2 equals V3, impedances Z1, Z2 and Z3 are chosen equal. This balance is obtained by means of a small DC motor energized by the amplified error signal to vary V3 by setting potentiometer P1 according to the error signal amplitude and polarity. The potentiometer P2 (ganged with P1) connections are crossed to directly read V4 equals V3 equals V1 + V2.

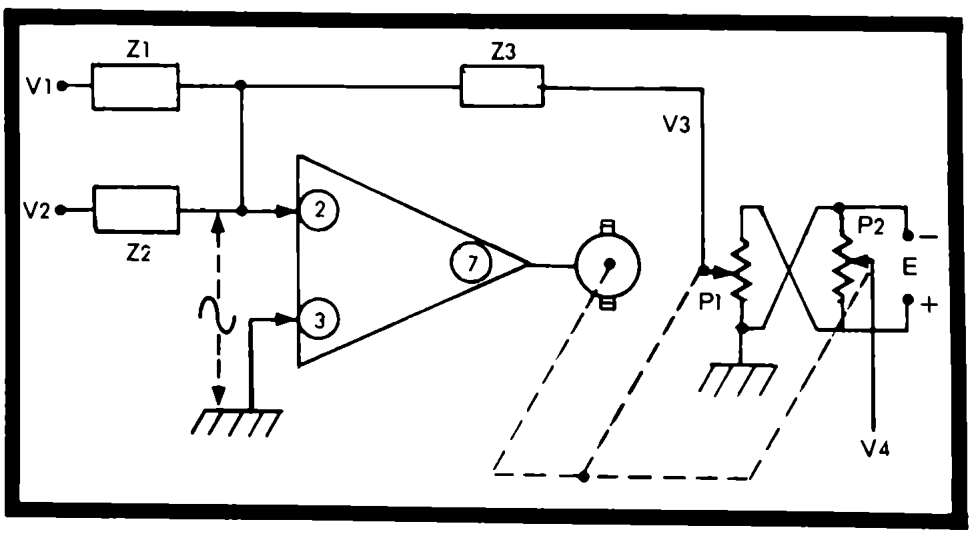


Fig. 5-55. In this circuit, analog voltage addition is accomplished by a servomechanism.

The two inverting and not-inverting inputs of an op-amp allow for conveniently carrying out **subtractions** by applying V1 and V2 to inputs 2 and 3 through series impedances Z1 and Z2; parallel impedance Z3 restores the input symmetry regarding inverse feedback impedance Z (Fig. 5-56). Supposing four equal impedances, V0 equals V1 - V2.

In the foregoing, speaking of impedances, the general case was considered. In the case of a DC analog computer, this reference evidently means resistances; in AC equipment, however, it may not necessarily mean resistors only; there can be capacitors and, though rarely, inductors. Signal polarity, then, is simulated by the in-phase or out-of-phase condition of the wave. On the other hand, if impedance equality facilitates an understanding of the mechanism of computing, it severely limits the operating possibilities of the computer.

The mathematical function of **integration** of a variable voltage is simulated by connecting a capacitor (C) in the feedback loop and a resistor R in series with the inverting input (Fig. 5-57). Since the amplifier input impedance is virtually infinite, the output voltage is given by:

$$v_o = \frac{1}{RC} \int_o^z v_i\ (t)\ dt$$

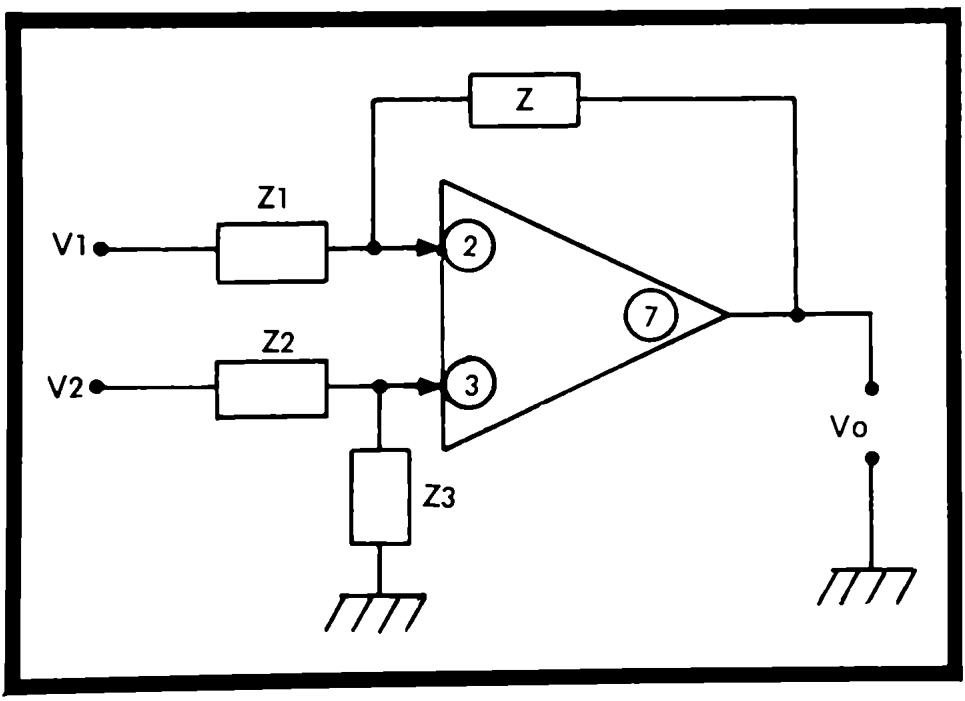


Fig. 5-56. Voltage subtraction circuit using the symmetry feature of an opamp.

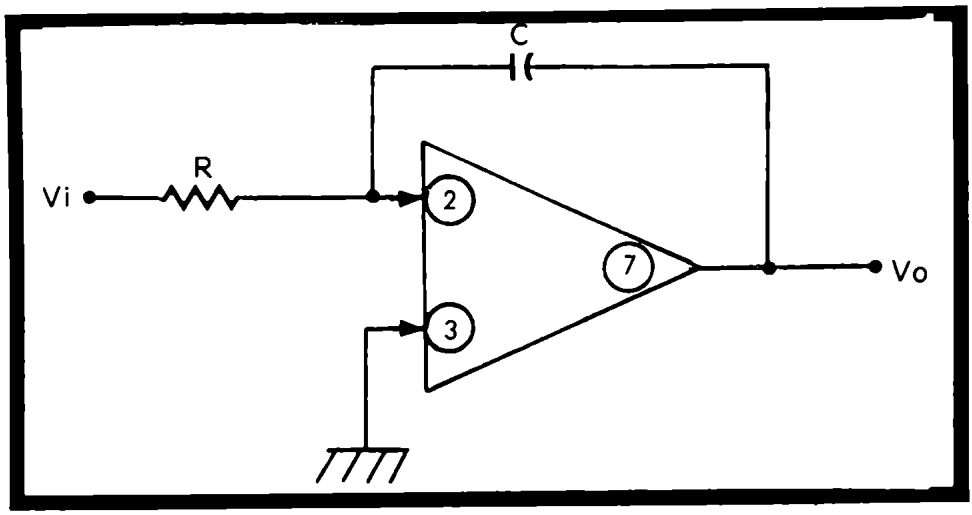


Fig. 5-57. With capacitive inverse feedback and a series input resistance, this opamp is wired to achieve analog integration.

with an integration time constant T equals RC.

Transposing R and C results in a **differentiating** circuit (Fig. 5-58) whose output voltage is:

$$v_o = RC \frac{dv_1}{dt}$$

the differentiation time constant again being T equals RC.

A **phase lag** can be simulated by shunting a capacitor C across feedback resistor R (Fig. 5-59) and a **phase lead** by

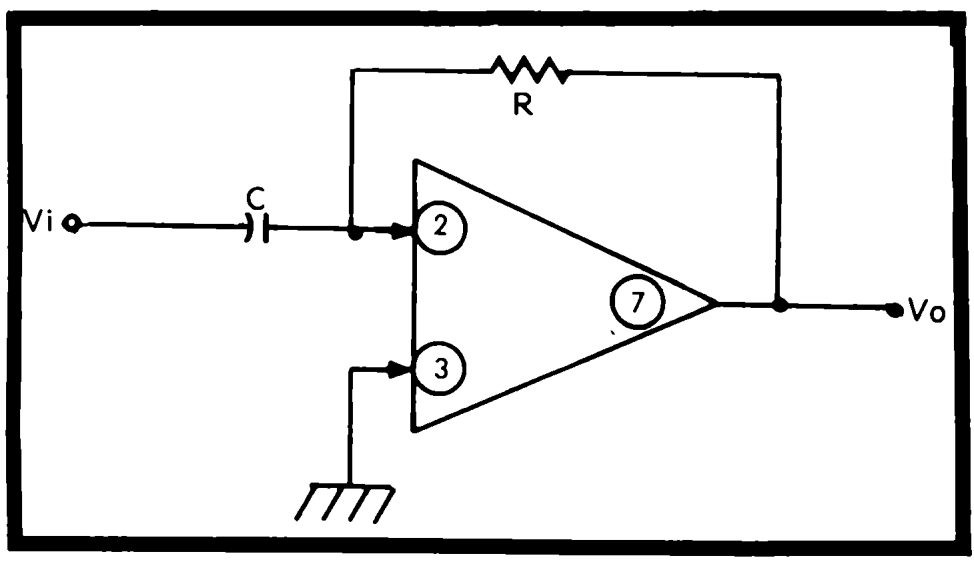


Fig. 5-58. Analog differentiating circuit. Compare with the integration circuit in Fig. 5-57.

connecting C across input series resistor R1 (Fig. 5-60). With some insignificant accessories, the opamp, thus, can be enabled to simulate a number of physical phenomena.

The Hall effect considered for transducer applications is discussed in Chapter 2 and the multiplier function of the Hall wattmeter is illustrated in Fig. 2-66. A **Hall-effect voltage multiplier** is represented schematically in Fig. 5-61. Input voltage, V1, is converted into a proportionate current, i_B, which excites the magnetic-field generating winding M by high-gain amplifier A1, using an inverse feedback by a common input-output impedance to stabilize the system transfer admittance. Similarly, input voltage V2 is converted into a proportionate Hall current, i, applied to Hall plate H by amplifier A2. The Hall voltage, VH, developing across H, is applied to differential amplifier A3. Since:

$$i_B = \frac{V_1}{R_1} \qquad \text{and} \qquad i = \frac{V_2}{R_2}$$

$$v_o = nV_H = \frac{n\,k}{R_1 R_2} \cdot V_1 V_2$$

Resistors R1 and R2 can be made variable to be used as input voltage coefficients.

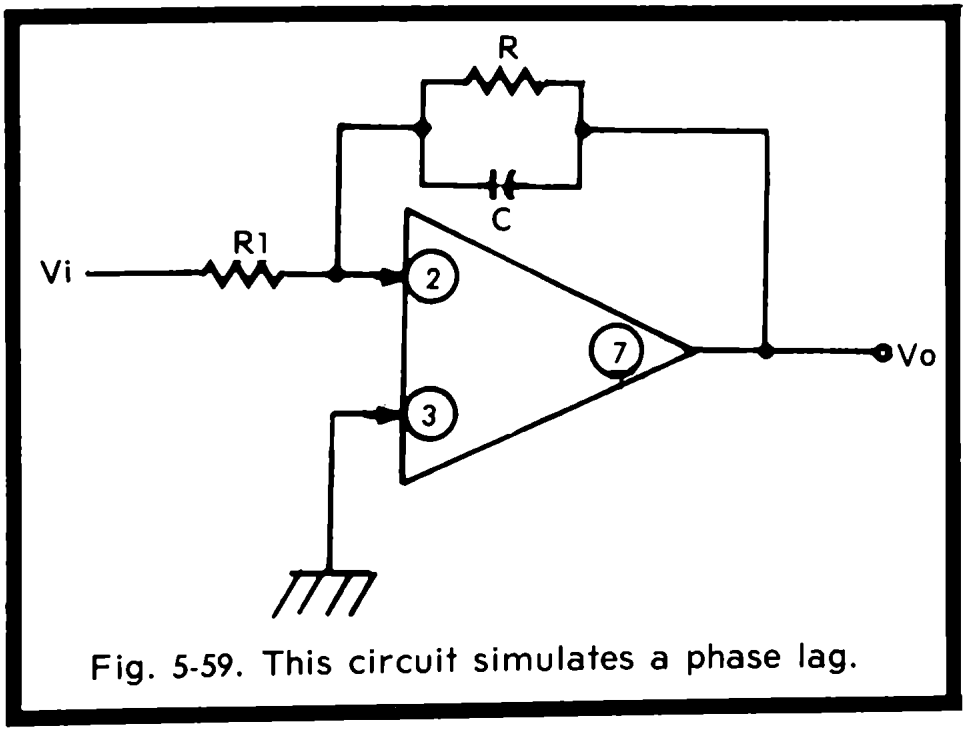


Fig. 5-59. This circuit simulates a phase lag.

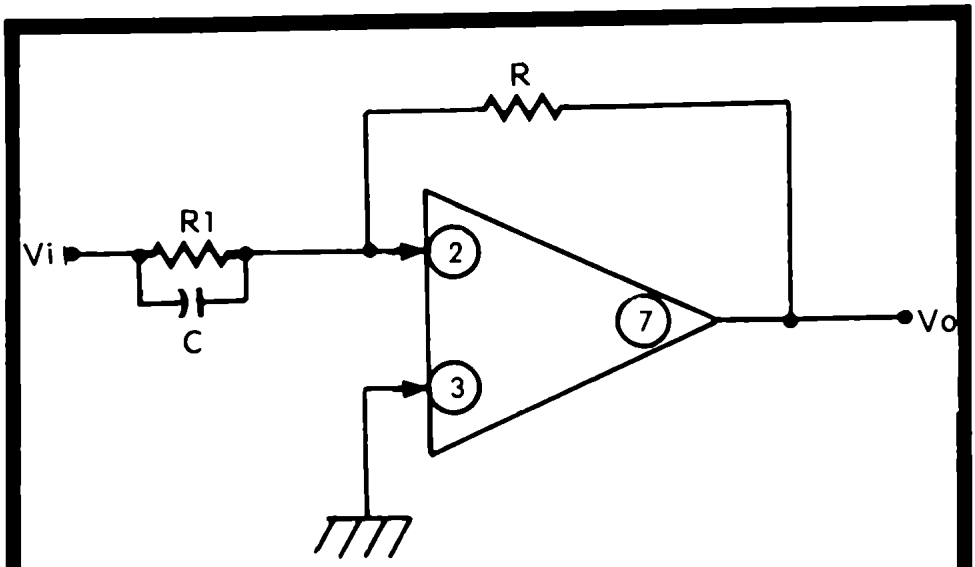


Fig. 5-60. This circuit simulates a phase lead.

AUTOMATIC CONTROL OF COMPLEX PROCESSES OR AUTOMATION

The extensive line of automatic machine and process controls presently available allows for automatically carrying out all machining and processing operations with the utmost speed and accuracy. Even machine feeding with raw materials and semi-finished items often is accomplished automatically. Only maintainance and repair still require the presence of skilled technicians, because automatic equipment is expensive and down-time is costly.

It takes more than merely automating individual machines to arrive at full-scale **automation** or entirely automatic production from raw material supply to finished product output. The entire production process has to be integrated by an overall control system taking into account all factors influencing product quality, quantity output and cost, supervising the raw material inventories, making the corrections a situation may need and signaling any condition requiring human intervention. This "super-control," actually a digital computer, monitors the operation of all individual machines down the line by suitable transducer signals, keeps the raw material inventory record, supervises the conveyor system carrying the workpiece from one machine station to the next, and tests and counts the finished products. The same computer is conveniently used for cost accounting on a daily basis, thus providing management with badly needed in-

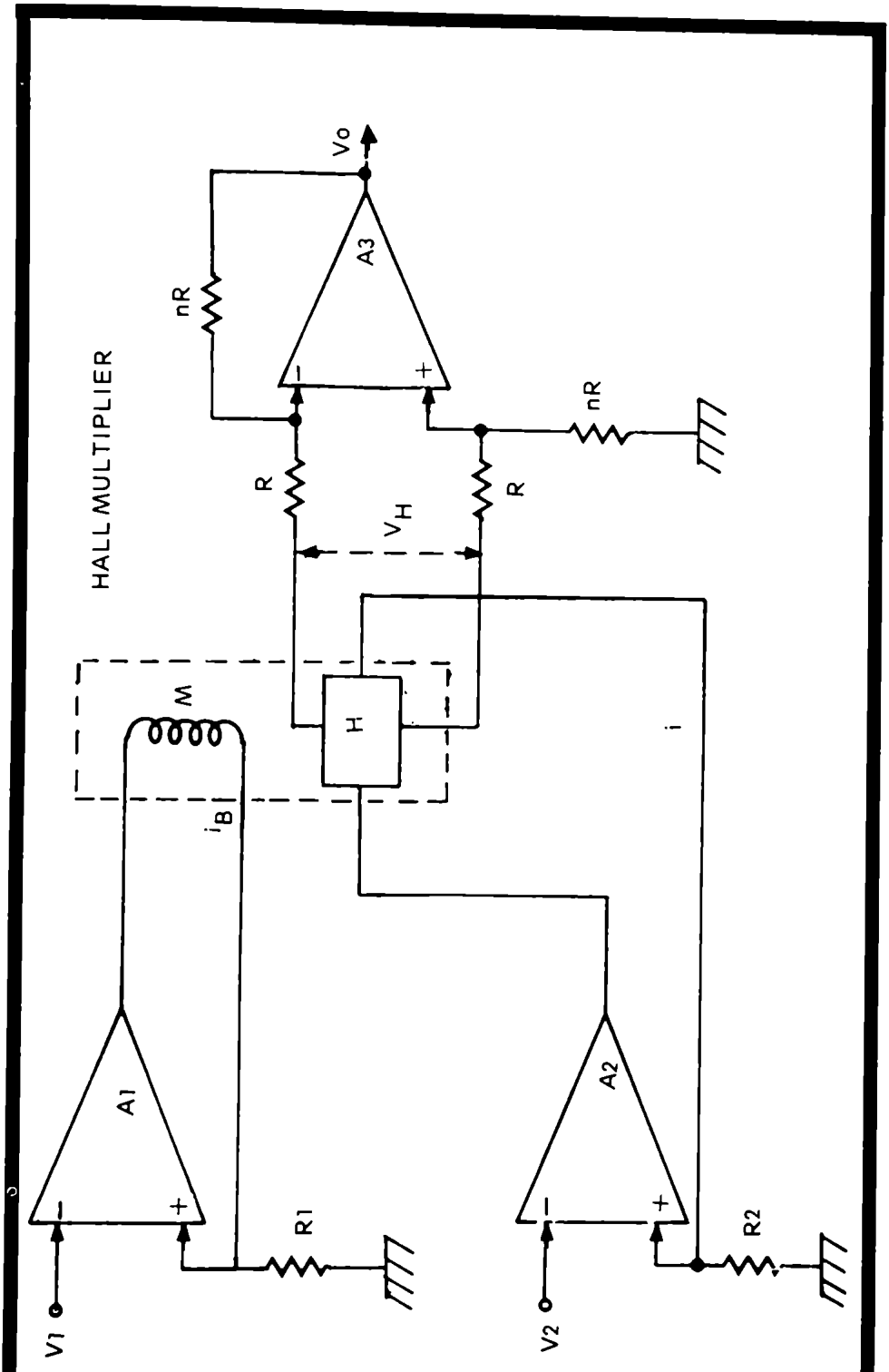
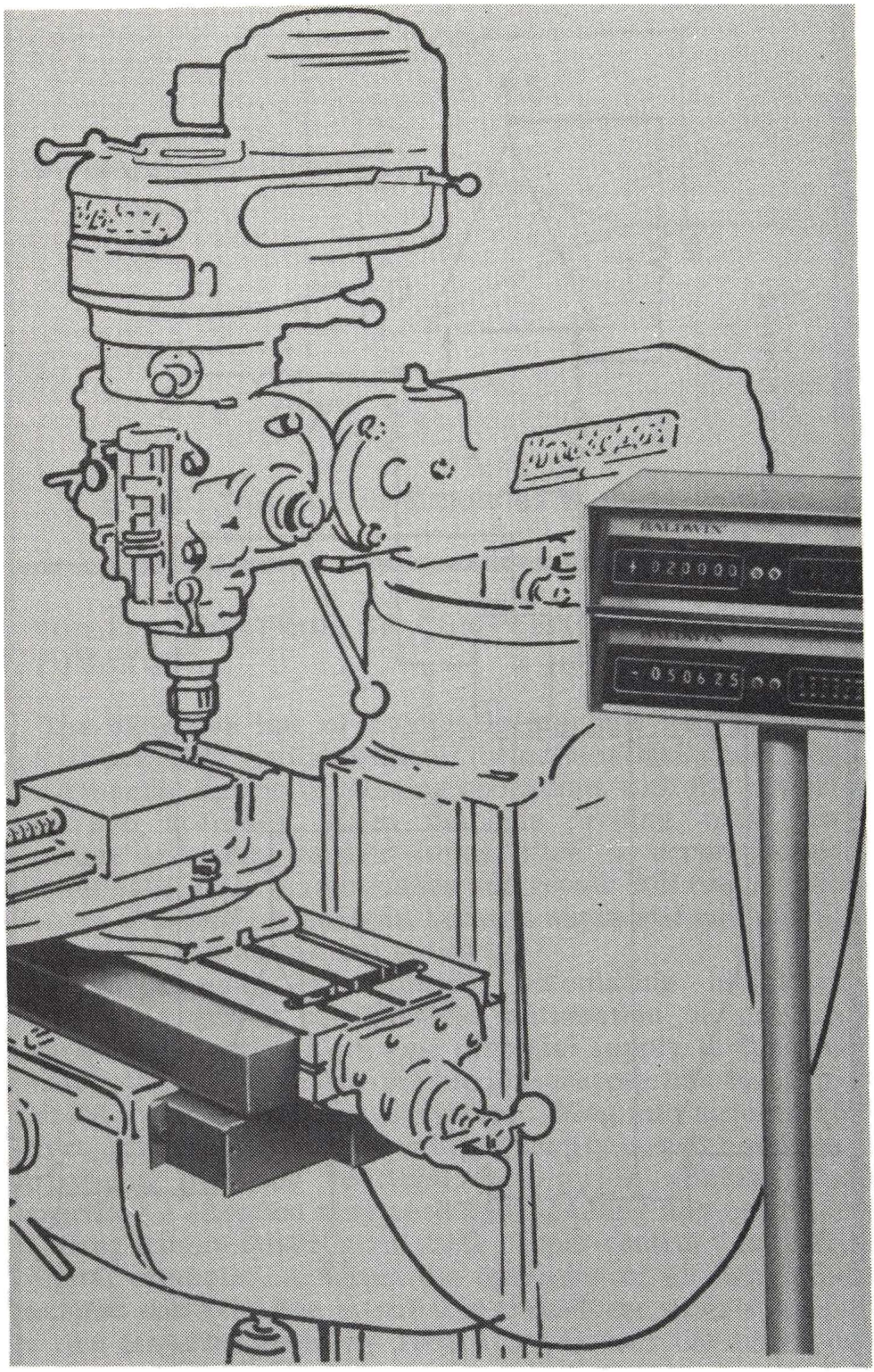


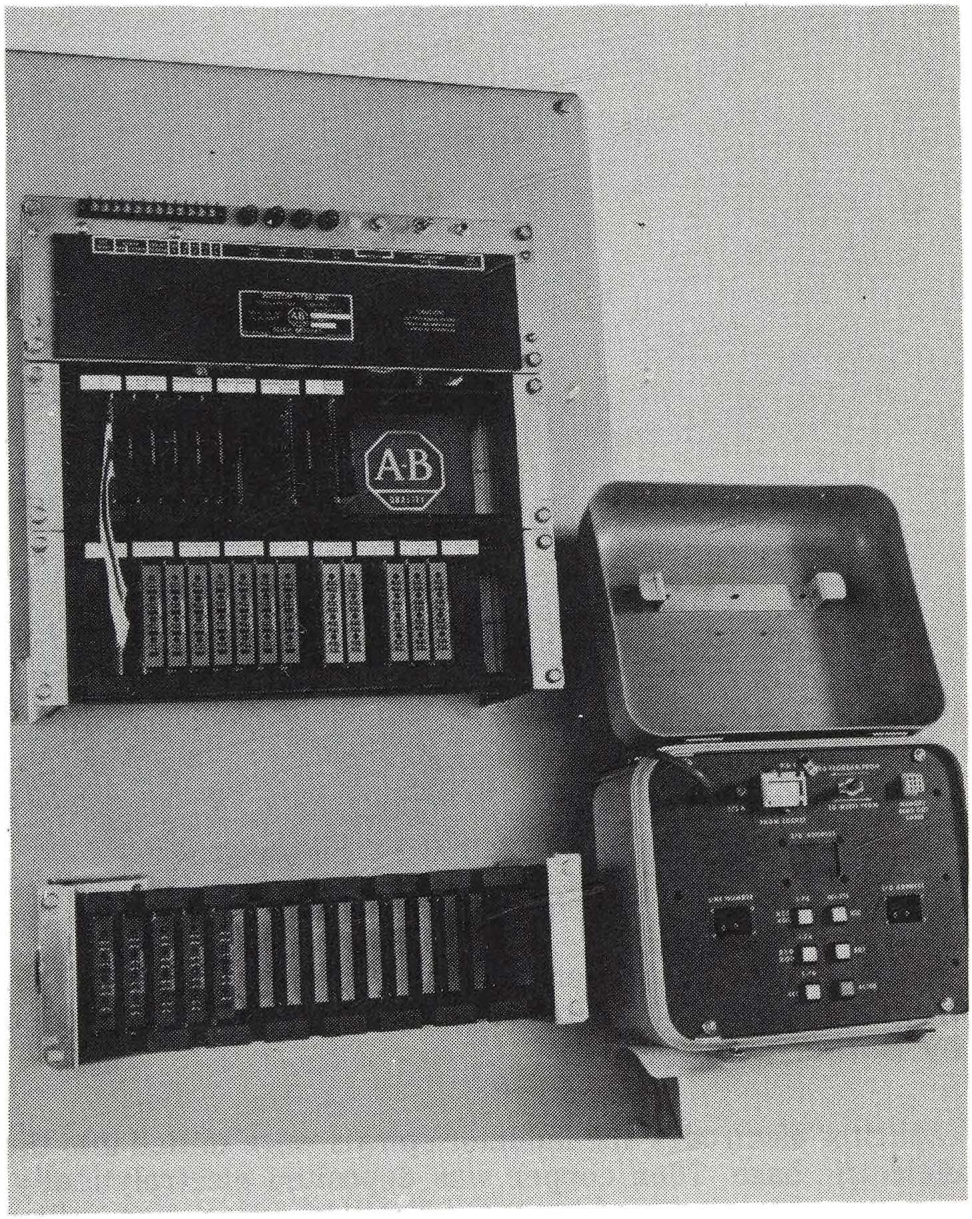
Fig. 5-61. Hall-effect voltage multiplier circuit.



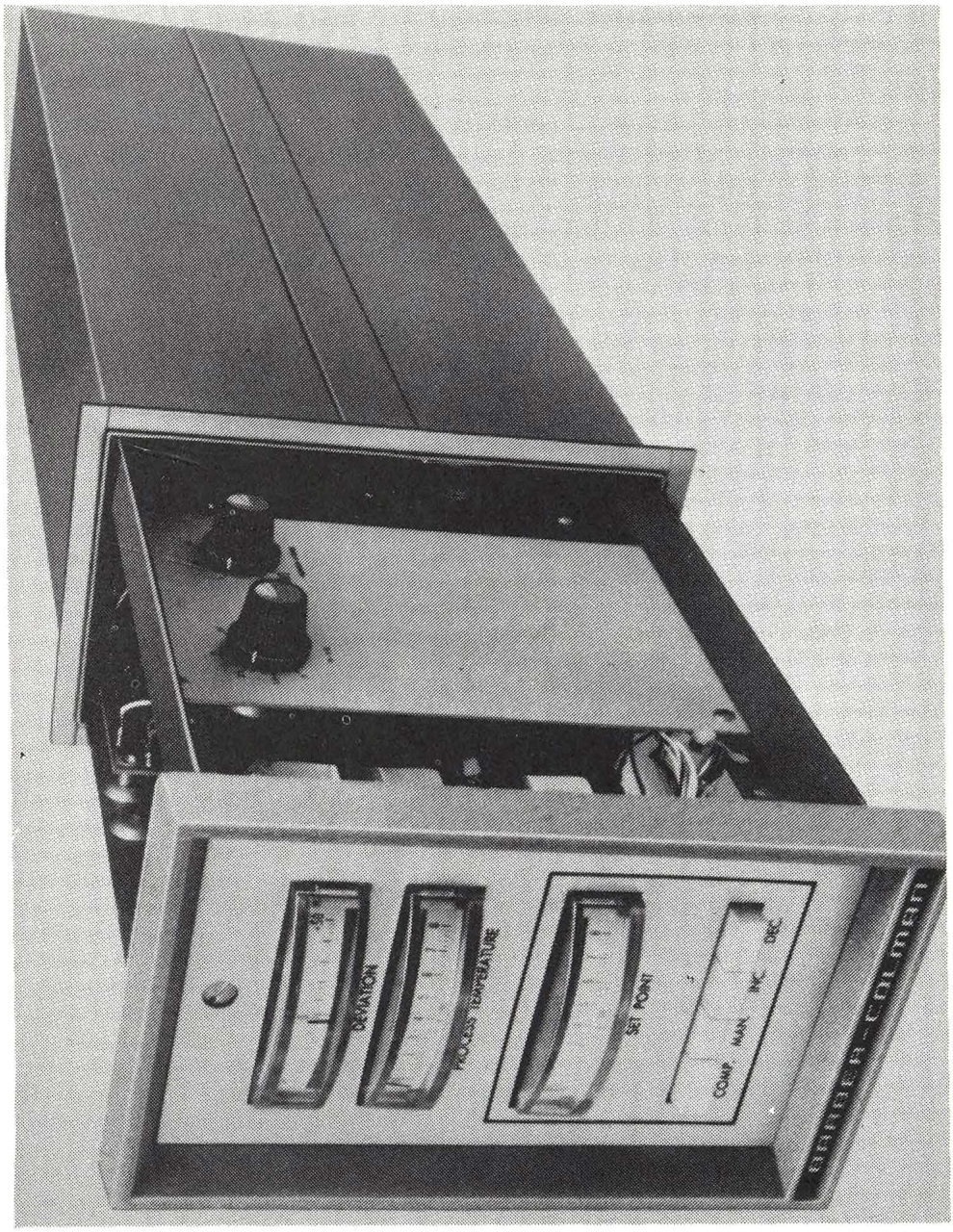
Dual axis, digital linear position readout system. The X and Y encoders on the table indicate the true positional data to 0.001 inch. Manufactured by Baldwin Electronics.

formation; and as labor and energy costs are factors in product cost evaluation, the computer also keeps an account of power spent as well as of hours worked, and can even print pay checks.

An early example of automation is the experimental setup for radio receiver manufacture developed by Sargrove in England in the early '50s. The somewhat simplified radio used a phenolic base plate introduced at one end of a huge machine. Since this was prior to printed circuits, the conductor paths,



Programmable matrix controller designed to control relays, motor starters, solenoid valves and indicates. Manufactured by Allen-Bradley Co.



Computer set controller designed to indicate deviation from a set point and process temperature. Manufactured by Barber-Colman Co.

capacitor electrodes and inductor spirals were milled into the dielectric base. Then copper was deposited electrolytically, and the excess copper on the nonmilled areas was milled away. Every machining step was followed by a dimensional check, and the built-in capacitors and inductors were measured and adjusted to specified values by copper removal.

The working radio chassis that came off this assembly line required only minor operations to yield a complete radio. With modern printed circuits and integrated electronics readily available, radios and other gear can be assembled a lot easier now.

Full production line automation calls for tremendous investments that can be justified only by mass production. On the other hand, modifications of such an integrated production system are expensive, so products prone to early obsolescence are unsuitable for automation. Semiconductor manufacturers, for one oppose full-scale automation of any production that may be obsoleted by continuous technological innovations long before the corresponding investment is written off. Since Detroit has made it a habit of obsoleting every year the former year's car, there is no point in automating car body building. On the other hand, a good, proven motor is assembled for years virtually unchanged on the chassis of changing cars and, therefore, qualified for automation. The chemical and oil industries process large quantities of products for years without any process changes; so their plants are extensively automated. The few people left are merely control and maintainance technicians.

CHAPTER 6

Electronic Heating, Welding & Machining

Conventional electronic methods are used concurrently for numerous production operations such as heating, welding and machining. In vacuum or controlled-atmosphere processing, melting and welding of refractory materials and working with very hard and brittle matter, electronics quite often offers the only methods practicable, especially in the production of delicate and comparatively costly components such as vacuum tubes, semiconductor devices and integrated circuits.

HIGH-FREQUENCY HEATING

An LC circuit tuned to the output of a high-frequency generator G (Fig. 6-1) sustains two kinds of losses—**dielectric** losses in the space separating the plates of capacitor C, and losses due to **eddy currents** and **magnetic hysteresis** in the magnetic field of inductor L. In communications, such losses must be kept as low as possible by suitable design and choice of materials, for they deteriorate circuit quality (expressed by Q factor) and overall performance.

However, since lost energy always manifests itself as heat, these RF losses can be used for heating various materials for soldering, welding and other industrial processes. Losses generally increase with frequency, but the choice of the optimum working frequency depends on the nature and the particular condition of the material to be processed. The energy is supplied by an RF generator very similar to a CW communications transmitter, although there is no provision for modulation or power supply filtering to eliminate hum. There are industrial RF generators rated from less than 1 KW to more than 100 KW, according to the particular application. For all but the lowest powers, transmitter-type vacuum tubes are used. This is one of the very few applications where they have not yet been superseded by solid-state devices because of their competitive cost, invulnerability to transients and high power capability.

Introducing the material to be processed in an LC circuit obviously detunes the circuit and change in the material's

characteristics during processing adds further detuning. Since the system efficiency is best on tune, continuous retuning by an automatic tuning control is necessary. Furthermore, such a powerful transmitter is liable to generate strong RF interference for nearby radio, television and communication equipment. Entrusted with the guardianship of communications, the Federal Communications Commission (FCC) has set strict regulations for radio interference by industrial generators, contained in Part 18, "Industrial, Medical and Scientific Uses of Radio."

According to the FCC rules, radiation must be carefully suppressed so that a signal of no more than 10 microvolts per meter can be detected 1 mile from the equipment, and the same limitation applies at a distance of 50 feet between the power line and the equipment to prevent the power line from acting as an antenna for the interfering energy. Since perfect shielding of rooms housing RF machinery would be extremely expensive for high powers and so small a radiation limit, the FCC set aside four frequency bands for operation of industrial heating equipment without license where radiation is not objectionable: 13.553 to 13.556, 26.960 to 27.280, 40.660 to 40.700 and 2,400 to 2,500 MHz. But even in the assigned frequency bands, harmonic radiation must be carefully suppressed so as not to exceed the 10 microvolt per meter limit at a 1-mile distance.

To comply with such stringent rules, elaborate and costly arrangements may be needed, such as shielding of the entire RF processing room. If a meshed wire screen is used for shielding, the mesh must be small enough, for any opening

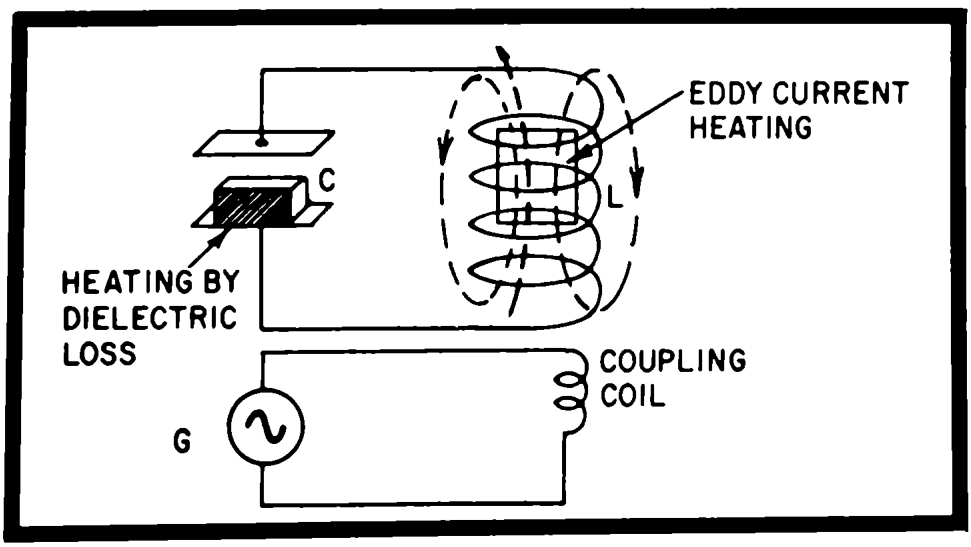


Fig. 6-1. RF heating is achieved by either dielectric losses or eddy currents.

greater than a quarter-wavelength would be just like no shielding at all. Doors must be covered with the shielding material, carefully bonded on the hinge side by flexible braid, and connected on the opposing side by a pressure contact. Power and any other lines entering the shielded room are to be carefully bypassed and screened to prevent them from radiating through the power network.

Depending on the capacitive or inductive nature of the RF losses, there are **dielectric** and **inductive** heating systems. Except for frequency and power requirements, both use the same generator type, but their applications differ.

Dielectric Heating

If a nonconductive material is placed between the plates of a capacitor (C) energized by a suitable RF generator, the rapidly reversing electric field traversing the material violently agitates its molecules, and this agitation generates heat. Unlike conventional methods where radiated heat penetrates from the outside, dielectric heating thus provides a uniform temperature rise from within throughout the heated substance (assuming the material is homogeneous). This avoids local overheating, and the amount of generated heat is readily predictable. Since the process stops when the resonant circuit is de-energized the heating time is under positive control.

The operating frequency depends on the dielectric losses of the particular substance and ranges from 2 to 100. Because these losses increase with frequency, low-loss dielectrics, such as most plastics, require a higher frequency. On the other hand, wood gluing, as in plywood manufacturing, is done at frequencies from 2 to 15 MHz because the water contained makes for higher dielectric losses. The necessary power usually ranges from 1 to 30 watts per cubic inch, but occasionally may run as high as 600 watts per cubic inch.

An important application of dielectric heating is the preheating of thermosetting plastics prior to molding. The granular or powdered material pressed into tablets is placed for some minutes between the horizontal plates composing the capacitor of a dielectric oven. The rated power ranges from 1 to 30 KW, and the operating frequency ranges from 10 to 40 MHz. Preheating results in faster and better curing in less molding time (resulting in an increased productivity of the molding press) and in a better product homogeneity. Similarly, in the rubber industry, dielectric heating is used to dry and cure a thick mass of latex uniformly throughout from

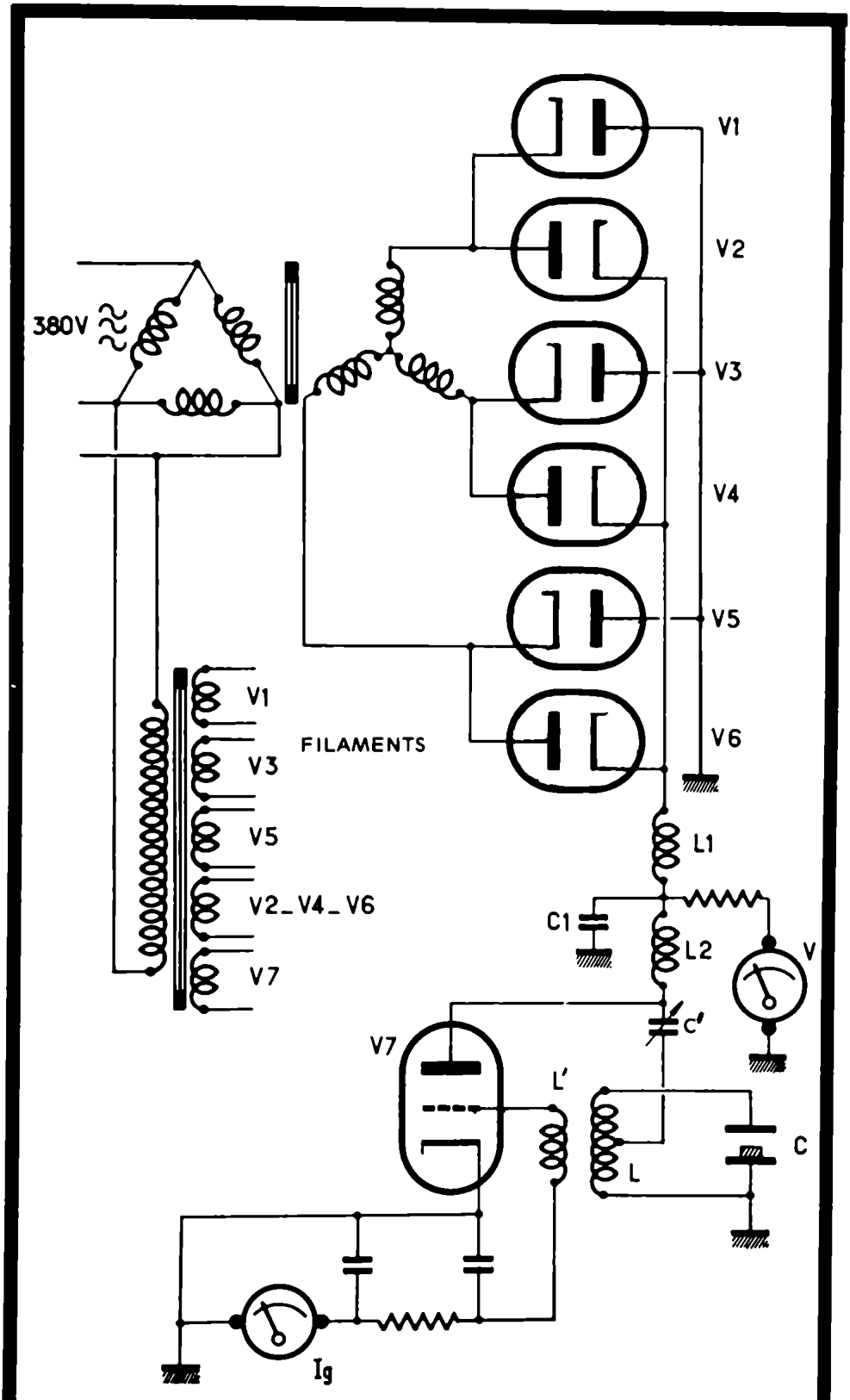


Fig. 6-2. Schematic of a dielectric heating generator.

the center in only four minutes, instead of the 40 minutes required by the former vapor processing.

Thermoplastic sheet welding is another important application of dielectric heating. Widely used in the manufacture of plastic raincoats, handbags, pouches and packaging materials, this "electronic sewing machine" employes an RF generator of some hundreds of watts of rated power working at 20 to 100 MHz. The capacitor is formed by narrow polished metal electrodes positioned on both sides of the plastic sheets to be welded and energized upon compression. The heat released within the material by its dielectric losses produces a fused joint, while the electrodes themselves remain cool and do not stick.

Still other applications include drying various fabrics, wool, cotton, as well as grain and food products. Medical diathermic equipment also uses dielectric heating; the material processed in this application is human tissue. In the UHF ranges, heat is generated, as far as is known, by the combined dielectric and inductive losses within the food (meat) subjected to the UHF field. Such ranges operate on a frequency of about 2500 MHz (wavelength 10 cm) with 1 to 10 KW of power generated by a magnetron or klystron oscillator. The LC circuit is replaced by a tuned cavity fitted with a door to introduce food to be cooked, or to be reheated if it is frozen. As the heat is released from within, the cooking or reheating takes only minutes. (For the same reason, such food lacks the browned crust that so many of us are fond of.

The simplified diagram of a dielectric heating generator is shown in Fig. 6-2. Its power is supplied by a 380v 3-phase delta-star connected transformer and six mercury-vapor rectifier tubes called **phanotrons** (Chapter 8). Rectified B+ is 2 to 10 KV, and even more. Filter L1C1 is meant only to prevent RF leakage, not to reduce hum.

The oscillator circuit uses a power triode, V7, in a tuned-tank arrangement and dielectric heating capacitor (C) forms a part of resonant circuit LC, coupled to the high-voltage anode of V7 by adjustable coupling capacitor C'. Oscillation is excited by feedback coil L' in the grid circuit. Grid current meter Ig reads oscillation intensity. A voltmeter (V) is generally provided to indicate the presence and actual magnitude of the anode voltage; these two instruments do for checking operating conditions, also.

In other generators, heating capacitor C is part of a resonant circuit, LC, coupled to the tuned oscillator tank, L'C', by a low-impedance line (Fig. 6-3). This arrangement has several advantages: the generator can be remotely located in regard to the working electrodes; the power applied can be

proportioned by varying the inductive coupling; and a dielectric breakdown between electrodes will not damage the generator.

In dielectric heating generators, a thyristor-controlled antiflash device is often provided to interrupt the B+ supply as soon as a flashover occurs between electrodes. Instant interruption is essential, for such a disruptive discharge would char the plastic instead of welding it and burn the polished electrode surfaces.

INDUCTION HEATING

Induction heating depends on the hysteresis losses (in ferromagnetic materials only) and eddy current losses or short-circuit currents induced in a conducting mass subjected to an AC field. The metal ring in Fig. 6-4 is traversed by the AC flux of coil L. Hence, it will behave like a shorted transformer turn, and, provided it is prevented from moving out of the field, it can only convert the induced current into heat (Joule effect). The phenomenon remains unchanged if the ring is replaced by a plain disc, or even a conductive mass of any shape. On the other hand, a thin plate placed along the flux lines is only flux-traversed within its thickness, and, therefore, will be heated very little by the eddy currents.

Induction heating primarily concerns metals, but, generally speaking it applied to conducting materials. While

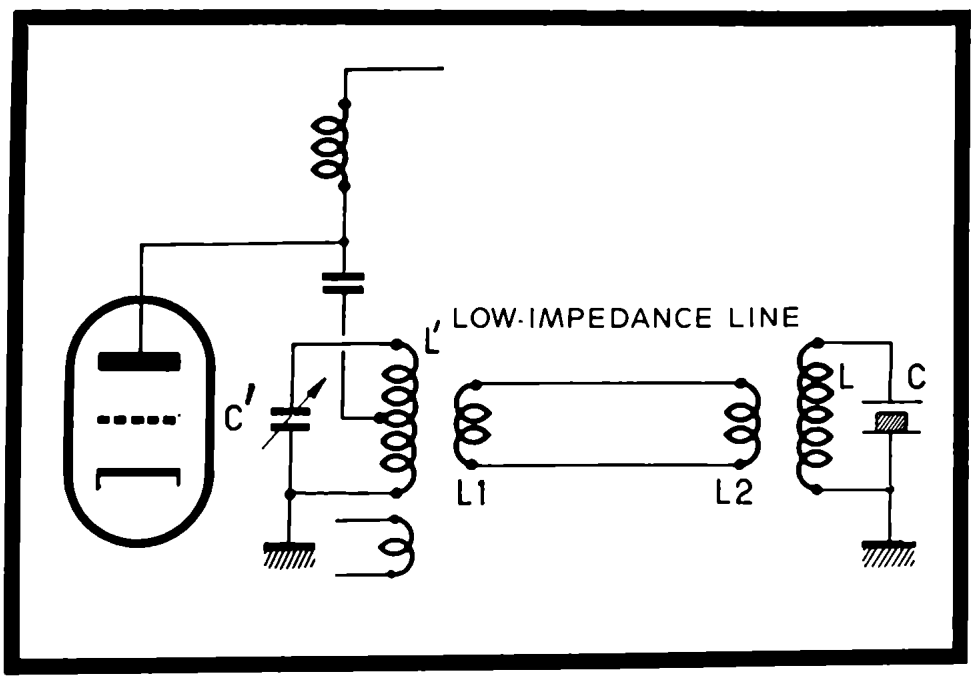


Fig. 6-3. Schematic showing a low-impedance line coupling the generator and load.

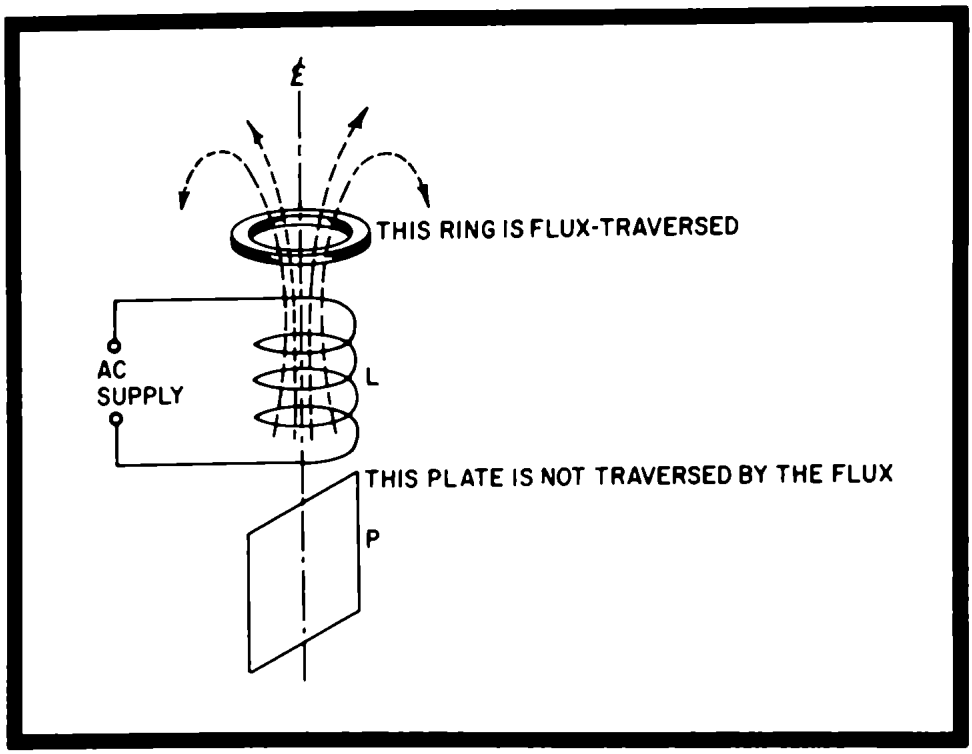


Fig. 6-4. In induction heating, eddy-current heat is induced in flux-traversed metal masses.

glass, for instance, is known as a good dielectric, it actually becomes a conductor in the melted state, so it can be first melted by conventional means and then processed in an induction oven. Insulating and semiconducting materials can be induction heated using conducting crucibles made of carbon or metal. In ferromagnetic metals, such as iron and steel, the heat generated by hysteresis will reinforce the action of eddy currents, at least up to their Curie point where all magnetism vanishes.

The higher the frequency, the more the induced field forces the electrons to the outer surface of the flux-traversed mass, so that the depth of current penetration decreases when the frequency increases. This phenomenon is called **skin effect**. Opposing dielectric heating, RF induction heating is not homogenous (at least at higher frequencies). This property is useful for **surface hardening** of gear wheels calling for a hard tooth surface for wear resistance as well as a soft, ductile core. Using a specially shaped inductive heating coil (Fig. 6-5), each tooth or a portion comprising several teeth can be heated and tempered separately, while the other regions remain soft. The penetration depth of the heat is under control by a suitable choice of operating frequency and processing line.

The induction heating coil closely surrounding the workpiece to be processed is, of course, severely heated by the radiated heat (not to speak of Joule effect heating), and must be vigorously cooled. Therefore, such coils are often made of a high-conductivity copper tube which also conducts the cooling fluid. The induction coil also can be made of a suitably bent copper flat on which is hard-soldered the copper tubing containing the cooling fluid (Fig. 6-6). To prevent formation of calcareous deposits within the cooling tubes, the cooling water is softened or distilled water is circulated in a closed refrigerating circuit.

Such heating coils are designed for a number of applications and can be readily adapted to special contours; a heating loop may even be introduced into a bore hole for internal hardening. In the merry-go-round machines used for automatic vacuum tube processing, a coil surrounding the

Large induction generator being adjusted by a technician.

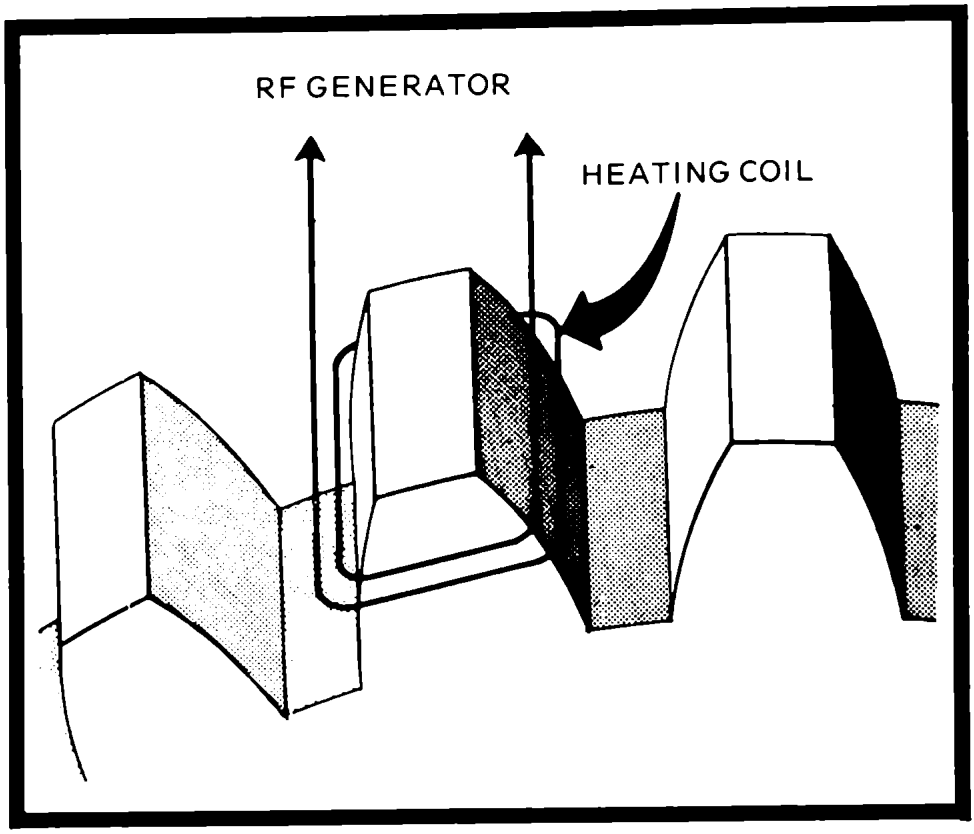


Fig. 6-5. Using a specially shaped coil, a single gear tooth can be hardened by induction heating.

glass envelope rapidly turns the electrodes enclosed red-hot to outgas them. And once the tube is sealed, another coil heats to evaporation the getter material laid down on a special closed-loop bracket forming a shorted turn.

Materials such as germanium and silicon that must be processed in vacuum or under a protective atmosphere are placed in a crucible made of a suitable material that is introduced in a quartz tube, which is sealed after providing the required atmosphere. The whole affair then is heated by a surrounding coil (Fig. 6-7). This method allows for the contamination-free processing essential to semiconductor and other materials. The floating-zone technique of semiconductor purification extensively used in production is practical only by induction heating. The ingot is made to slowly pass through a narrow heating coil, melting just the zone comprised within the turns; the remaining zones remain solid or are solidified again. The impurities float at the forefront of the molded zone.

RF induction heating generators mostly operate at frequencies from 0.2 to 2 MHz with from 10 to 100 KW of power. Soft and hard-soldering operations can be performed using preformed solder alloys fused in place.

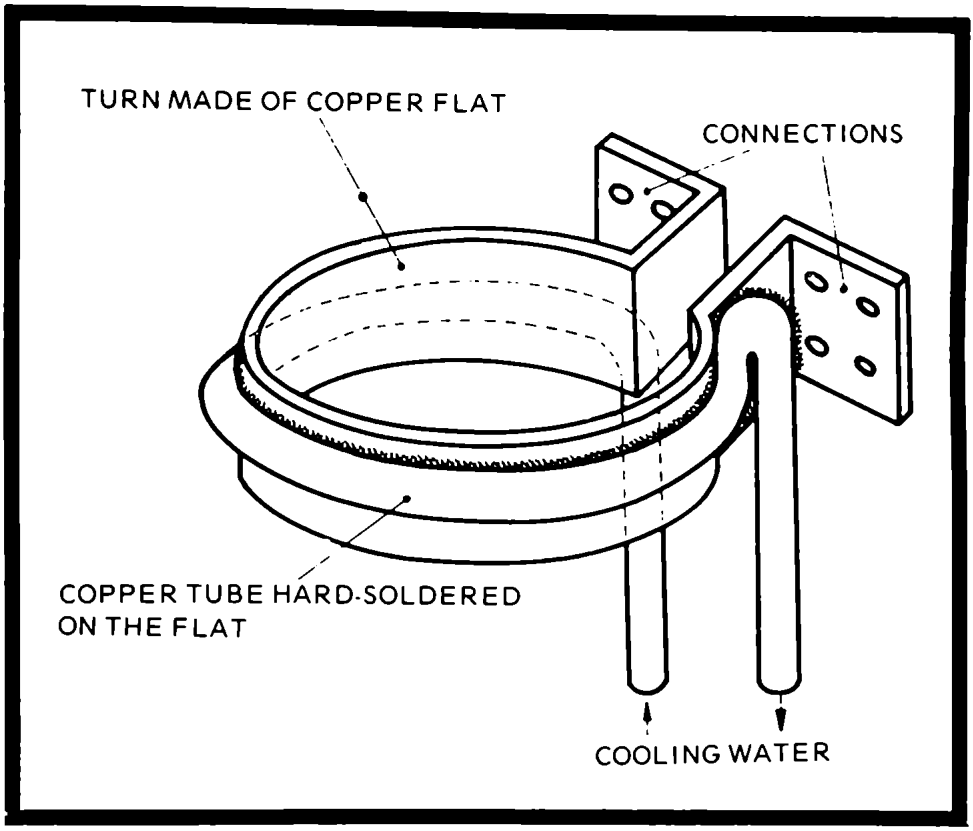


Fig. 6-6. Drawing of a 1-turn water-cooled induction heating coil.

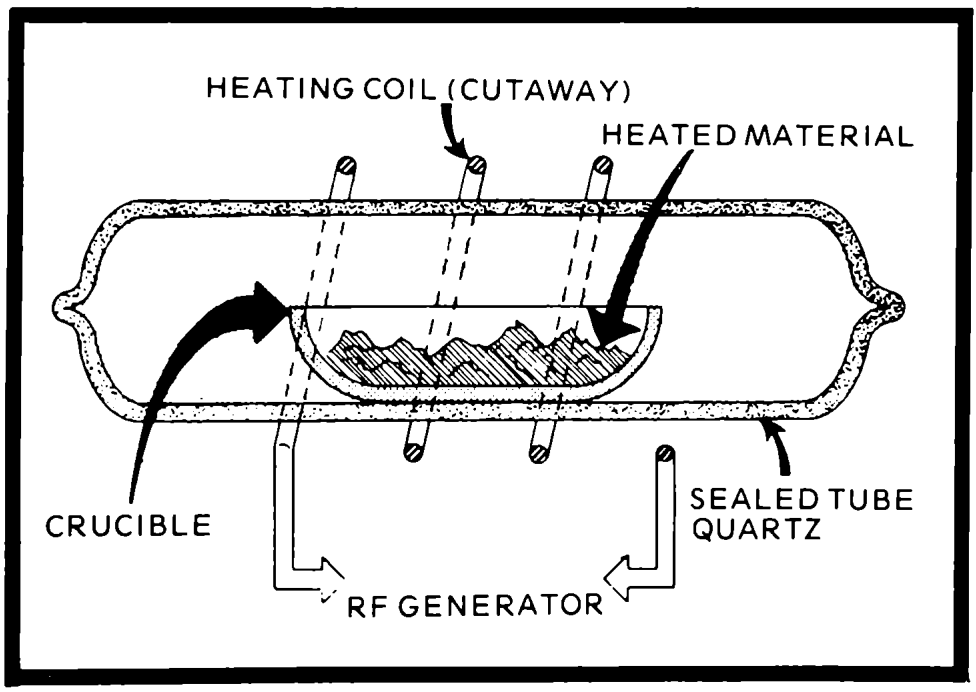


Fig. 6-7. Drawing showing inductively heated material processing in a protective atmosphere.

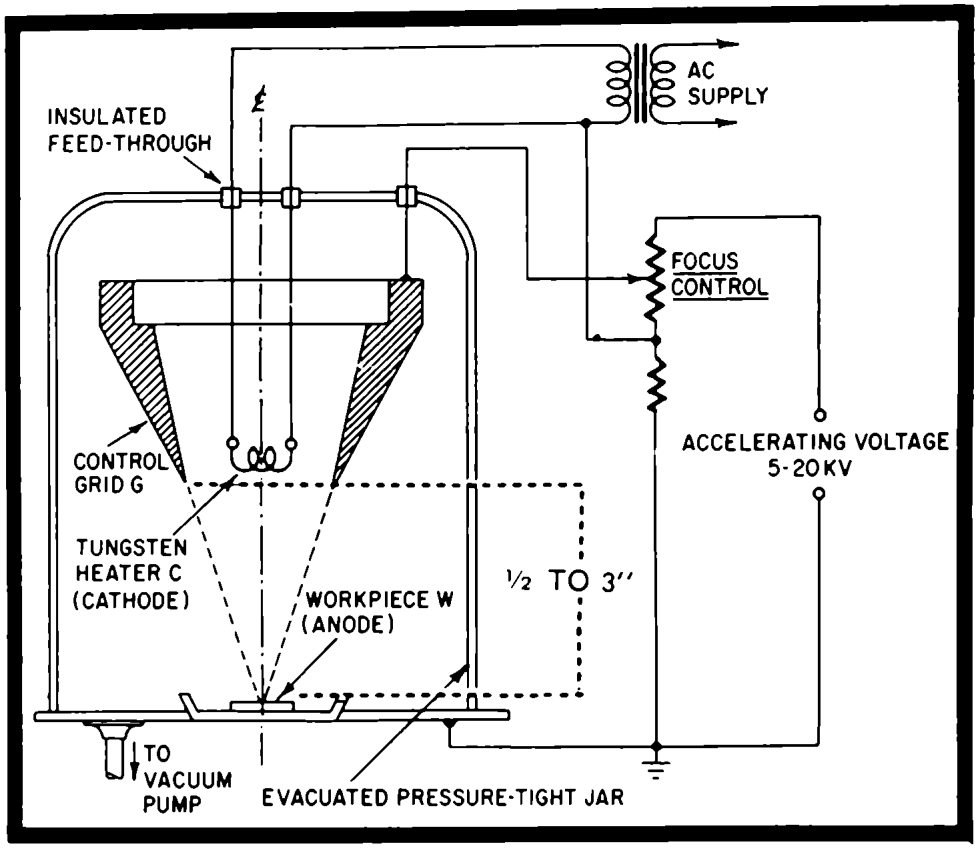


Fig. 6-8. Diagram of an electron-beam processing device.

ELECTRON-BEAM HEATING

This more recent process can be termed truly electronic, for the processing heat is produced by the kinetic energy of a fast electron beam hitting the work. Electron beam machines are available for melting, welding and machining conducting materials. Since the process takes place in an evacuated jar, this approach is particularly convenient for reactive metals such as zirconium and titanium. It also allows for processing refractory metals like tungsten and molybdenum, and high thermal-conductivity metals such as copper and silver that are difficult to process by conventional methods. Such sophisticated equipment is quite expensive, but aeronautic and aerospace industries can hardly do without it.

In an electron-beam machine, a tungsten heater-cathode emits electrons that are focussed by a negative-potential conical grid and strongly accelerated toward the work by a high-voltage supply (Fig. 6-8). The strongly concentrated beam hits the work in a very small impact spot (less than .01-inch diameter) where the considerable kinetic energy is dissipated, leading to immediate, intense heating of a minute

region of the work and quick fusion of the irradicated matter. Skillfully putting this heat to work, the target can be melted, welded or cut.

This phenomenon can only take place at a very low pressure (less than 2×10^{-4} mm Hg), so the whole device is enclosed in a pressure-tight jar, permanently connected to a vacuum pump. The need to introduce the materials to be processed into a vacuum chamber of obviously limited dimensions and to control the operations from the outside (possibly by a closed-circuit television system) evidently is a shortcoming, but vacuum processing is mandatory for some "exotic" materials and advantageous for a lot of others. And frequently, more convenient or less costly processing methods simply do not even exist.

The heat quantity Q (in calories per minute) dissipated in the impact spot is:

$$Q = 1.4(10^{-2}EI)$$

where E is the accelerating voltage and I the beam current (in amps). In the machine in Fig. 6-8, the work acts as an anode. Accelerating voltages of 5 to 30 KV are used with powers up to 5 KW; spot diameters range from 0.8 to 3 mm. For highest heat dissipation, the work is positioned in the focal point of the beam.

To obtain more heat concentration in a narrower fusion zone and to increase the penetration depth, magnetic focussing and higher acceleration voltages are used. With voltages of up to 100 KV and 1 KW of dissipated power, a beam of less than .01-inch diameter (0.25 mm) is obtained, allowing for cutting holes of intricate shapes. Such a machine is much like an electron microscope and costs about five times the price of the model above. Furthermore, the target hit by high-energy electrons emits dangerous X-rays and must be screened heavily to provide a safe working environment. Only qualified technicians can operate such a machine.

The **self-accelerated gun** combines the advantages of the (comparatively) low and high accelerating voltage systems. The gun in this system not only has a cathode and grid, but also a focussing, diaphragm-like anode and an accelerating anode with a center hole traversed by the pencil-shaped beam (Fig. 6-9). A focussing coil may be added as shown. Electrons now are accelerated within the gun before they leave by the hole. The workpiece is no more a part of the focussing system, and there is no electric field between the gun and workpiece. The spot diameter is from 0.8 to 3 mm, but can be as low as 0.25

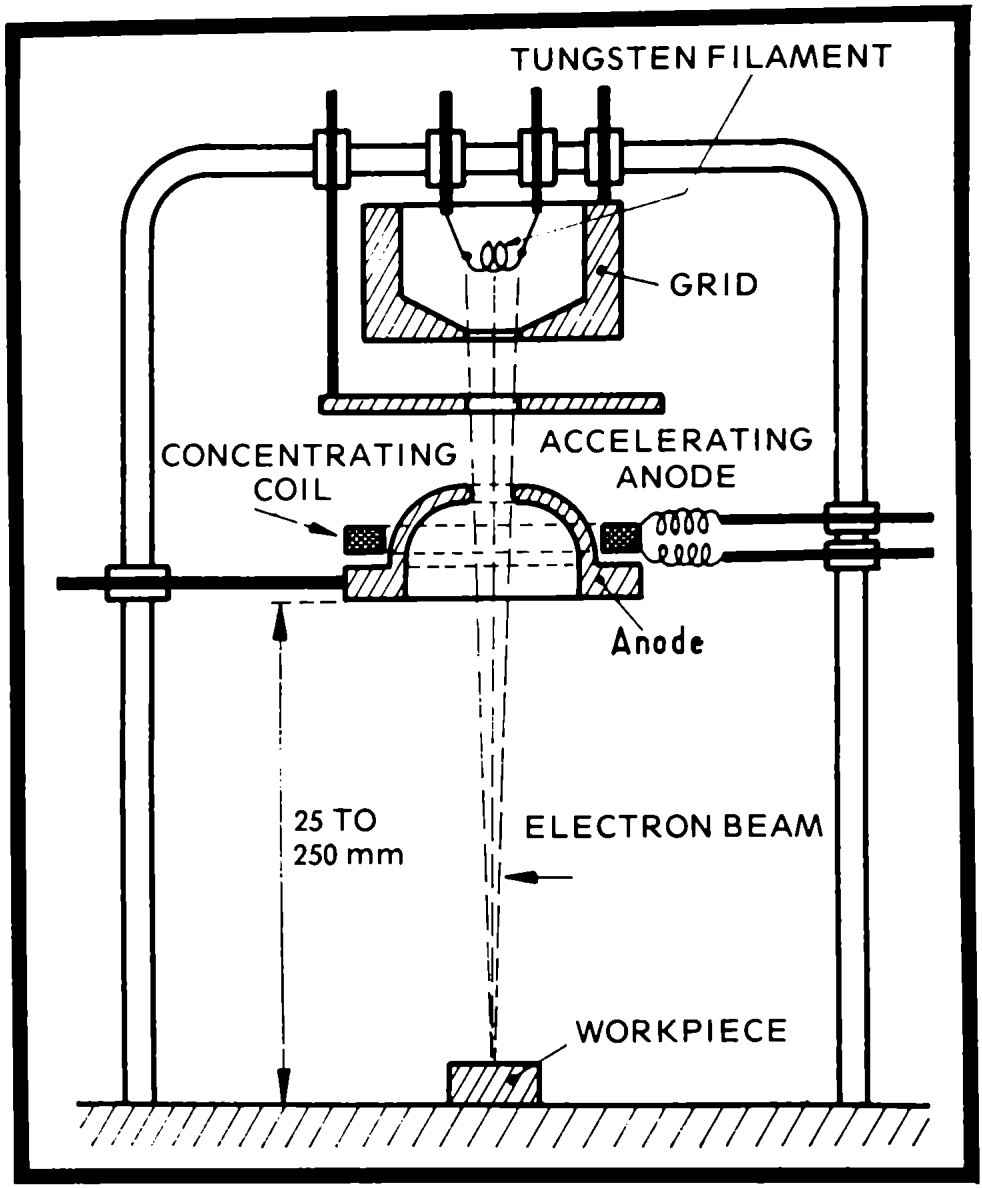


Fig. 6-9. Diagram of an improved self-accelerating gun electron-beam heating device.

mm with magnetic focussing. Accelerating voltages range from 10 to 30 KV with 5 to 10 KW of power.

ELECTRONIC RESISTANCE-WELDING CONTROLS

A resistance welder contains a transformer whose secondary is composed of a single turn of large-section copper conductor connected to heavy welding electrodes made of pure copper or copper alloy. A pedal or foot switch provides pressure to compress the parts to be welded between these electrodes (Fig. 6-10). This closes the transformer secondary

loop, and though only about 1 to 10 volts is induced in the single turn, the circuit resistance is very low, so that a high current builds up, instantly fusing together the workpieces. This is **spot welding**, extensively used in car-body and major-appliance assembly; by using wheel-shaped electrodes, sheet metals can be continuously seam-welded.

Resistance welding is just a variety of Joule-effect heating with a dissipated heat RI2, where I is the short-circuit working current. As virtually all the heat produced is to be released within the parts to be assembled, their resistance R must be high regarding the remainder of the circuit. Steel plates of much higher resistivity than copper are, therefore, easy to weld, while copper is difficult to join by this process, which evidently applies only to metals.

The welding transformer's primary current is controlled by a switching device called a **contactor**. Relays cannot handle such fast and accurately timed operating cycles nor the large currents involved; therefore, the switching is performed electronically by thyratrons and thyristors in small-power welders and by **ignitrons** (Chapter 8) in medium and large machines. A contactor is a mere switch and a **welding control** is necessary to operate it.

A typical high-power contactor is composed of two ignitrons, V1 and V2, anti-parallel connected for full-wave transformer operation (Fig. 6-11). Normally nonconductive, these ignitrons can be fired during the positive half-wave in

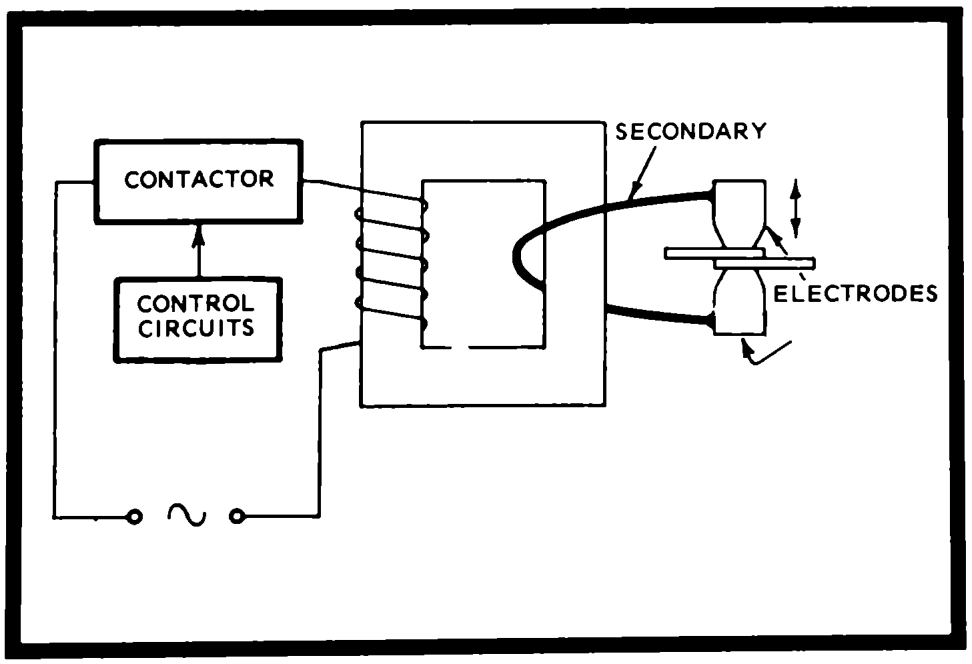


Fig. 6-10. In a resistance welder, heat is generated in a shorted transformer turn.

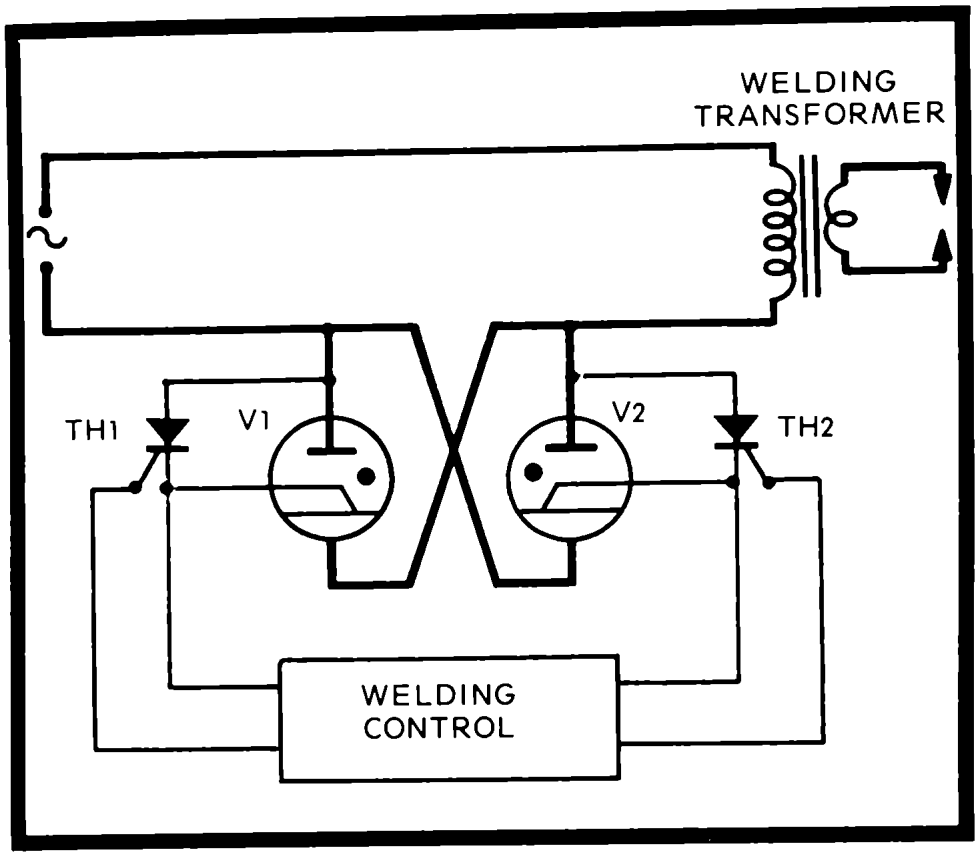


Fig. 6-11. Schematic of a ignitron welding control contactor.

each case by connecting the ignitor to the anode of the same tube through thyristors Th1 and Th2. The thyristors are fired by the welding control. As long as a thyristor is conducting, the associated ignitron fires at each positive half-cycle.

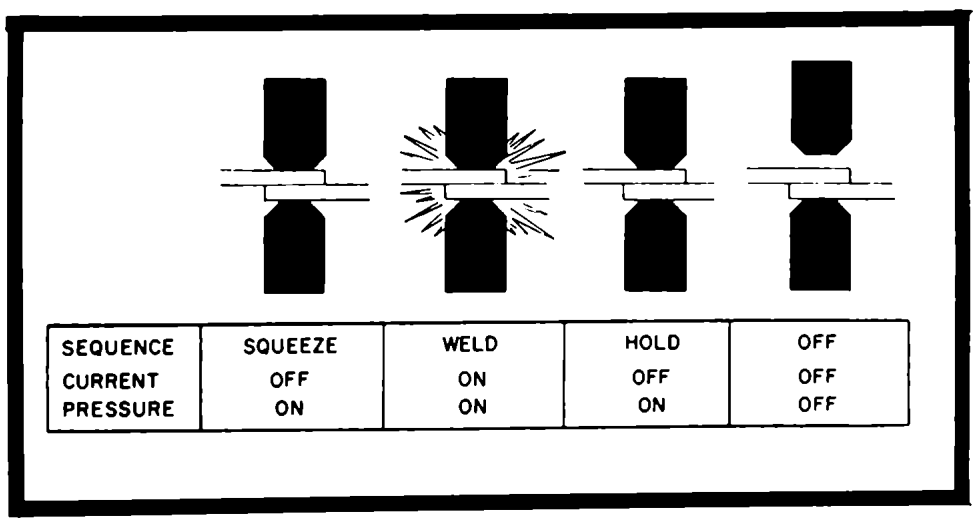
A resistance-welding process is performed in four timed steps according to Fig. 6-12: During the **squeeze time**, the workpieces are squeezed together by the electrodes to provide for good contact and to crush a possible oxide surface layer. The weld current passes during the **weld time**, locally fusing the squeezed materials. When the current ceases, the pressure is still maintained during the **hold time** to allow for solidification of the fused material. It is only in the **off time** that the pressure is released by raising the upper electrode to permit withdrawal of the welded work and the introduction of the next piece. All four steps need some **timing**, but the weld time must be set accurately, for it depends on the nature, surface condition and thickness of the materials to be joined. A welding timer provides adjustable time delays that are accurate as well as reproducible.

Timing generally is produced by charging or discharging a capacitor through a resistor or a constant-current device. The timer triggers a thyristor or similar device at a preset level, thereby ending the time delay. The time duration is then proportionate to the time constant, RC, or depends on the charging current and can be set by means of a potentiometer.

The timer in Fig. 6-13 (General Electric) is composed of a unijunction transistor (UJT) which generates a variable delay by means of R10 and C3, and a thyristor (TH) that energizes relay REL. The UJT circuit supply voltage is stabilized by zener diode DZ for accurate and reproducible time delays.

Upon connecting the voltage supply, TH conducts and energizes the relay, throwing contacts K1, K2 and K3 to a. To trigger the timing cycle, the start button is pressed; this shorts TH, which de-energizes, and with REL no longer powered, the three contact sets fall back to their rest position b. No longer shorted by K1, C3 charges through potentiometers P (calibration) and R10 (time adjust) until its voltage is sufficient to fire the UJT, which transmits a positive firing pulse to the gate of TH. The relay is powered again, and the contacts switch over to working position a to end the cycle. Contact set K3 closes the controlled circuit during the time interval generated.

Connecting the right-hand start-button connection to contact a of K2 prevents the beginning of a new cycle before the preceding one is completed, and thus from initiating a faulty timing cycle. In case of an emergency, the stop button



SEQUENCE	SQUEEZE	WELD	HOLD	OFF
CURRENT	OFF	ON	OFF	OFF
PRESSURE	ON	ON	ON	OFF

Fig. 6-12. Sketches illustrating the four steps of a resistance welding process.

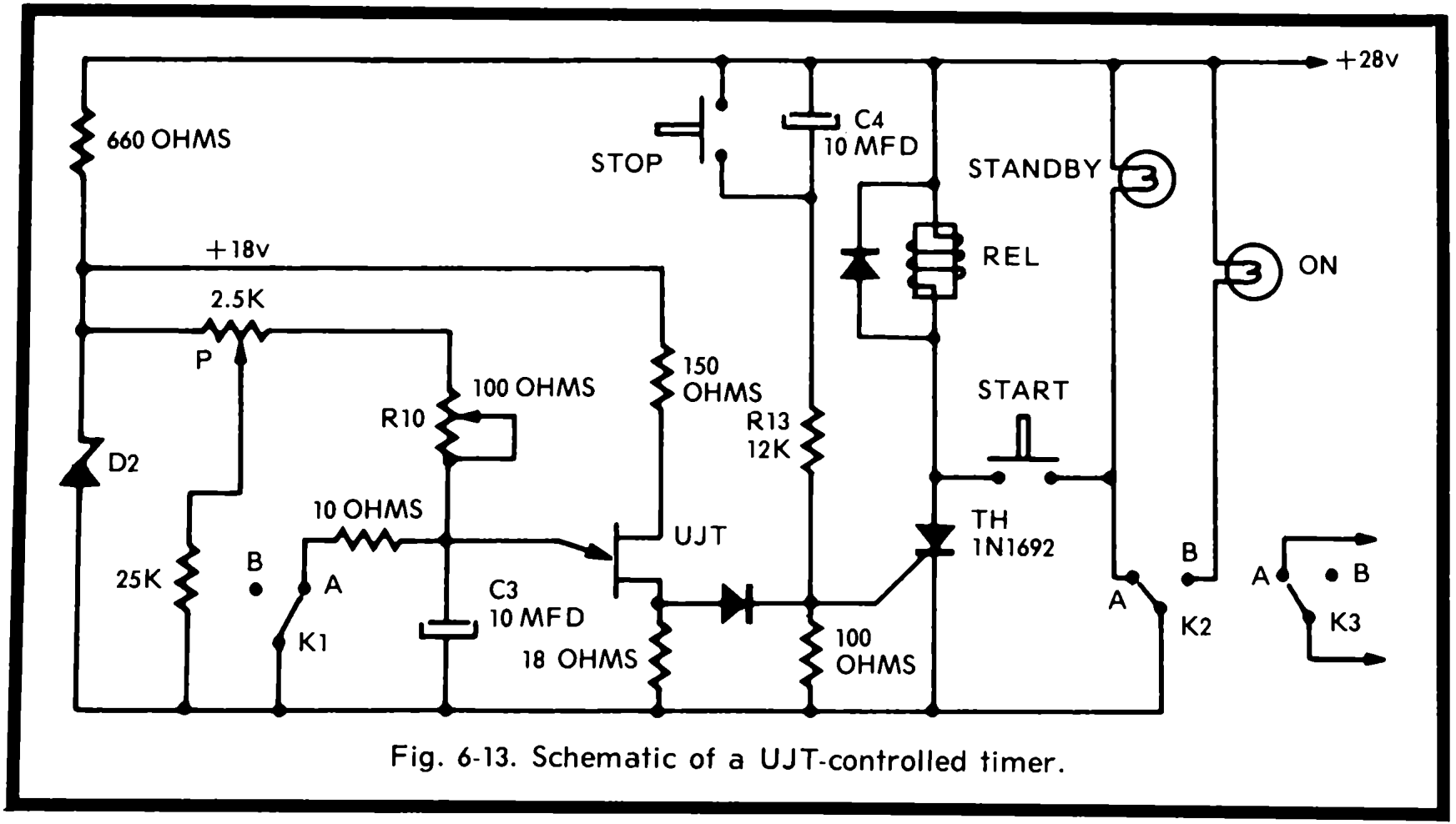


Fig. 6-13. Schematic of a UJT-controlled timer.

allows the cycle to be interrupted by passing a current through R13 to fire TH.

In addition to this "analog" circuit which allows the selection of any possible time interval there are **digital timers** that count a preset number of half-waves (discontinuous control). For 60-Hz mains operation, one or perhaps two counter decades generally will do.

SYNCHRONOUS LOW-POWER WELDING CONTROL

To avoid high power transient demands that arise when the welding transformer is connected at the crest of the AC supply voltage, the operation is preferably initiated at a low-voltage point of the cycle. Furthermore, if the welding always starts at the same point of the cycle, the results will be consistent and repeatable. It also is desirable that the welding operation always take an **even** number of cycles, starting with a positive and ending with a negative half-cycle, to avoid residual flux in the transformer core that could alter the initial conditions of the next weld.

All these conditions can be met by a **lag-tube leading-tube** circuit such as shown in Fig. 6-14, which illustrates a low-power welder using triode thyratrons V1 and V2 of adequate power, instead of ignitrons, and a small tetrode thyratron V3 for simple static timing. By means of pulse transformer T2, the lead tube (V1) is always fired first at a given point of the voltage cycle, adjusted once and for all by phasing resistor R1. The control grid of the lag tube (V2) is connected to feedback transformer T3, whose primary parallels welding transformer

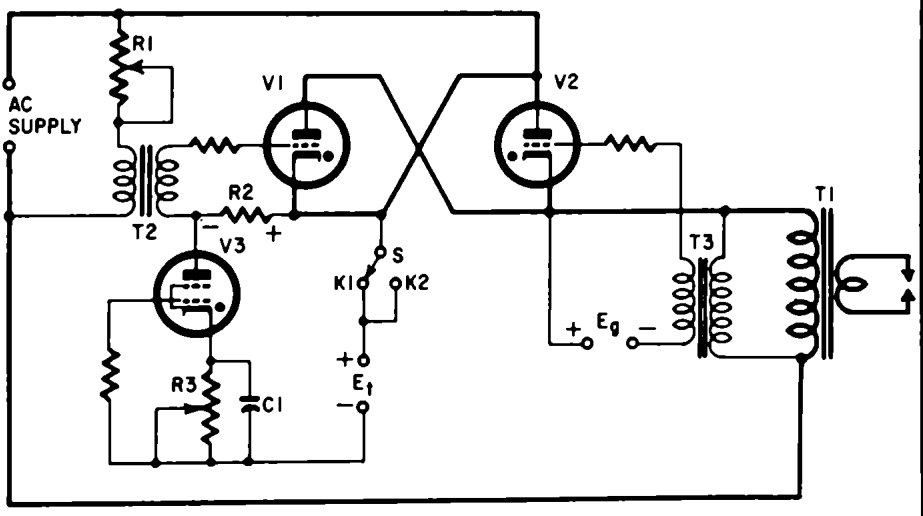


Fig. 6-14. Circuit of a synchronous welding control.

T1. Its bias voltage (Eg) prevents conduction unless V1 fires and passes a current through T3's primary, inducing in its secondary enough voltage to fire V2 on the next half-cycle. Thus, each time V1 fires (and only then), V2 will fire on the next half-cycle and welding current will always pass for an even number of cycles.

The timer is remarkable for its simplicity. A small thyratron (V3) fed by DC voltage supply Et fires as soon as the control is energized, and its current passing through resistor R3 develops a voltage drop that biases V1 sufficiently to prevent it from firing. Simultaneously, this current produces a voltage drop across cathode resistor R3 and charges capacitor C1. This strongly biases the control grid of V3, but as this tube conducts, its grid has no effect.

Throwing toggle switch S (of the break-before-make type), with contacts K1 and K2 tied together, momentarily interrupts V3's plate current, and on making contact again, its high cathode bias prevents the thyratron from striking. Since a voltage drop no longer develops across R2, V1 fires upon reaching the selected point on the cycle and V2 follows. Meanwhile, C1 discharges through R3, the bias voltage collapses. After a time delay, depending on R3 and C1, V3 fires again. This develops a new voltage drop across R2, which prevents, V1, and V2 by the same token, from firing.

In the controls described above, only the timing of the actual welding operation has been considered. Since the squeeze and hold times are not critical, they can be assigned fixed durations. (The off time obviously exists only in repetitive welding operations.) In a comprehensive welding control, the three or four timers are triggered sequentially. The squeeze is usually obtained pneumatically by a solenoid-operated valve that connects the cylinder of a pneumatic actuator to a compressed air supply. This valve closes during the off cycle, another valve allows the compressed air to flow out of the cylinder, and a spring or weight raises the upper electrode. The welding cycle generally is initiated by depressing a footswitch, leaving both hands free to guide the work. Often there is a repeat switch that restarts new welding cycles again and again, after completion of a definite off time, as long as the footpedal is depressed. If the repeat switch is open, the off time will last until it is closed.

ULTRASONIC PROCESSING AND MACHINING

Ultrasound is a form of vibratory energy; in fact, it is a very high-pitched and, therefore, an inaudible sound. Its frequency ranges from about 15 kHz to the highest megahertz,

and even gigahertz frequencies. (In fact, sonic frequencies often could be used too; but you simply couldn't stand the uproar near a high-power machine.) While low-power ultrasound can be directly produced by ultrasonic whistles and sirens, industrial applications requiring powers from, say, 100 watts to 10 KW use special high-frequency generators whose output is applied to transducers that convert the electric into vibratory energy, just like a speaker transforms audio voltages into sound.

Unlike electromagnetic waves propagating in air, and even in the vacuum of deep space, ultrasonic vibrations propagate best through dense media such as solids and liquids. Sound waves are severely attenuated by air and other gases and stopped altogether by a vacuum. To transmit ultrasonic energy to a workpiece, the transducer has to be directly in touch with it, or some liquid is used to transmit the vibrating energy. An airgap would absorb most of the useful energy. Ultrasonic processing or cutting, therefore, mostly takes place in a liquid-filled "dish."

Ultrasonic transducers are either piezoelectric or magneto-strictive. **Piezoelectric transducers** use quartz crystals, and nowadays piezoelectric ceramics, that vibrate longitudinally upon the application of an exciting AC voltage. The vibration amplitude, normally very small, can be enhanced by operating the transducer at its resonant frequency. This requires careful mechanical transducer design. In the device in Fig. 6-15, two piezoelectric blocks are bonded to three thick steel plates; the inner (insulated) plate is the "hot" electrode. Between the outer grounded electrodes, a double-amplitude vibration is available. The diagram of an ultrasonic (voltage) generator powering a 2-electrode

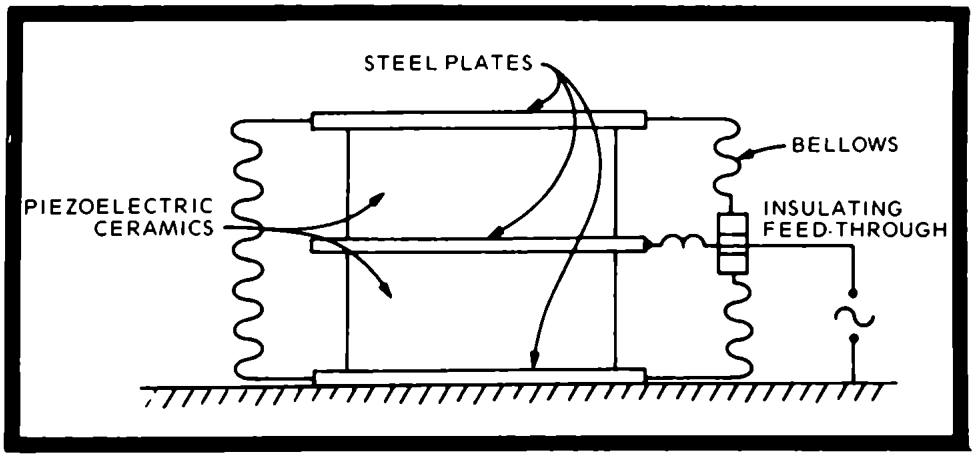


Fig. 6-15. Drawing of an ultrasonic transducer using two piezoelectric ceramic blocks held between steel plates which serve as electrodes.

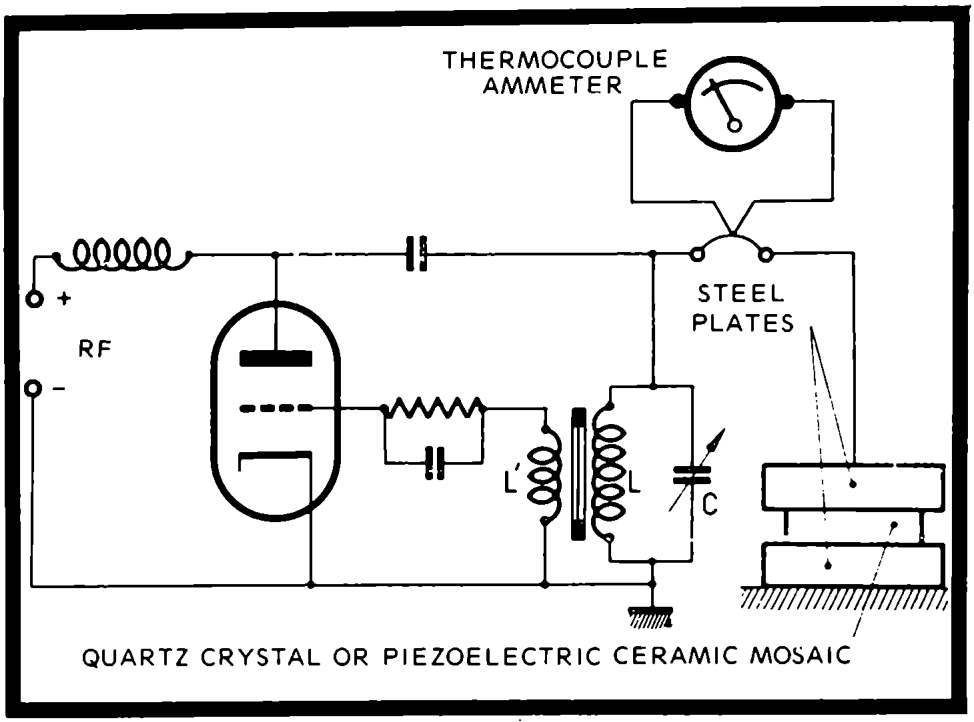


Fig. 6-16. Schematic of a piezoelectric ultrasonic transducer exciter.

piezoelectric transducer is shown in Fig. 6-16. The generator is carefully tuned to the transducer resonant frequency, for mismatch results in a considerably lowered energetic efficiency. Piezoelectric generators are operated at frequencies above 50 kHz, with ultrasonic powers of about 100w.

A nickel or iron-cobalt alloy contracts slightly on the application of a magnetic field. (Here, too, very small amplitudes are involved.) In **magnetostrictive** transducers, ultrasonic vibrations are generated by subjecting such a magnetostrictive core to an AC magnetic field by means of a high-frequency current-fed winding. To prevent eddy-current losses from converting a part of the applied energy into heat, the core is composed of thin, mutually insulated laminations of magnetostrictive material stacked and clasped to form a solid block. A relative amplitude variation of .003 percent or so is obtained; that is only 6 micrometers for an 8-inch core. By discerningly using resonant properties, much higher amplitudes are possible.

The diagram of relative contraction in terms of the magnetic field applied (Fig. 6-17) shows a nearly proportional rise at the beginning, followed by a gradual flattening. Therefore, there is no point in using magnetic fields as high as 100 oersteds. Practically, the core is biased by a DC magnetic field of 50 oersteds corresponding to operating point P, and an

AC field is superimposed to sweep P between M and N. As the instantaneous values of the AC field add and subtract from the DC field, the core alternately contracts and expands (an unbiased core can only contract).

The "works" of a magnetostrictive transducer are represented in Fig. 6-18. The core block bearing the exciting and biasing windings is firmly clamped at its center (nodal line xx') so that its upper and lower ends vibrate back and forth by successive contraction and elongation. Only one end, generally, is delivering ultrasonic vibratory energy by means of a coupling cone firmly bonded or welded to it.

ULTRASONIC CLEANING AND LIQUID PROCESSING

Fitting a suitable can to the upper end of such a transducer allows for ultrasonic vibration-treatment of liquids. By the action of the intense molecular agitation, the liquid releases gas bubbles the ascend to the surface. This phenomenon, called **cavitation**, is used for outgassing not only liquids, but molten metals as well. Some chemical reactions are accelerated by ultrasonic processing, for oxygen set free by cavitation has an especially active oxidizing property. And

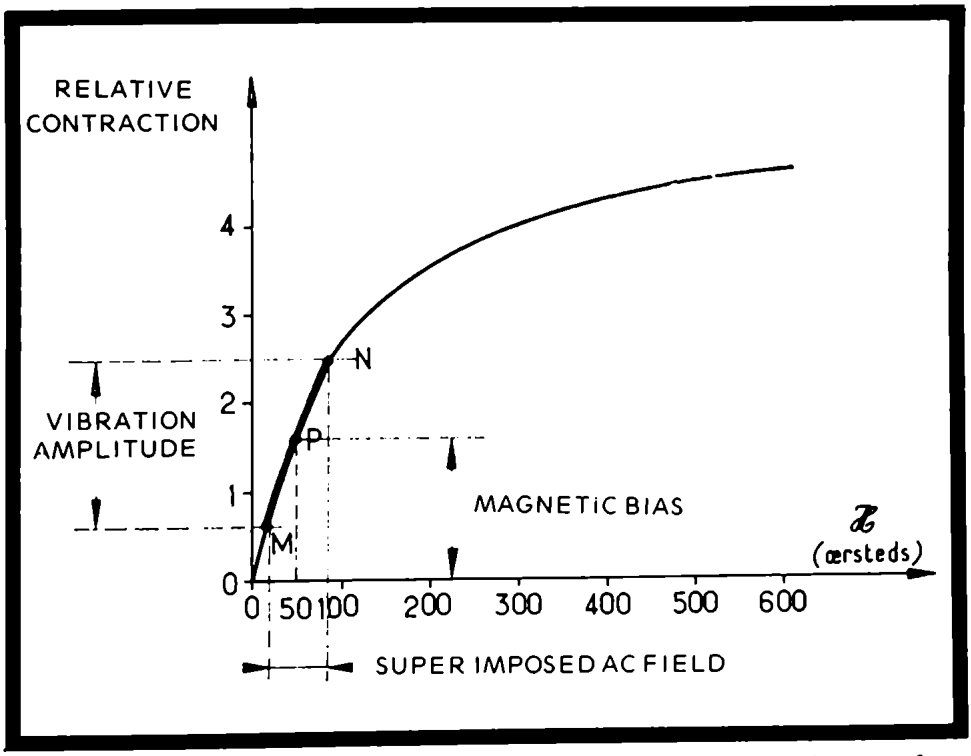


Fig. 6-17. Curve showing the relative contraction of a magnetostrictive material in terms of magnetic field strength.

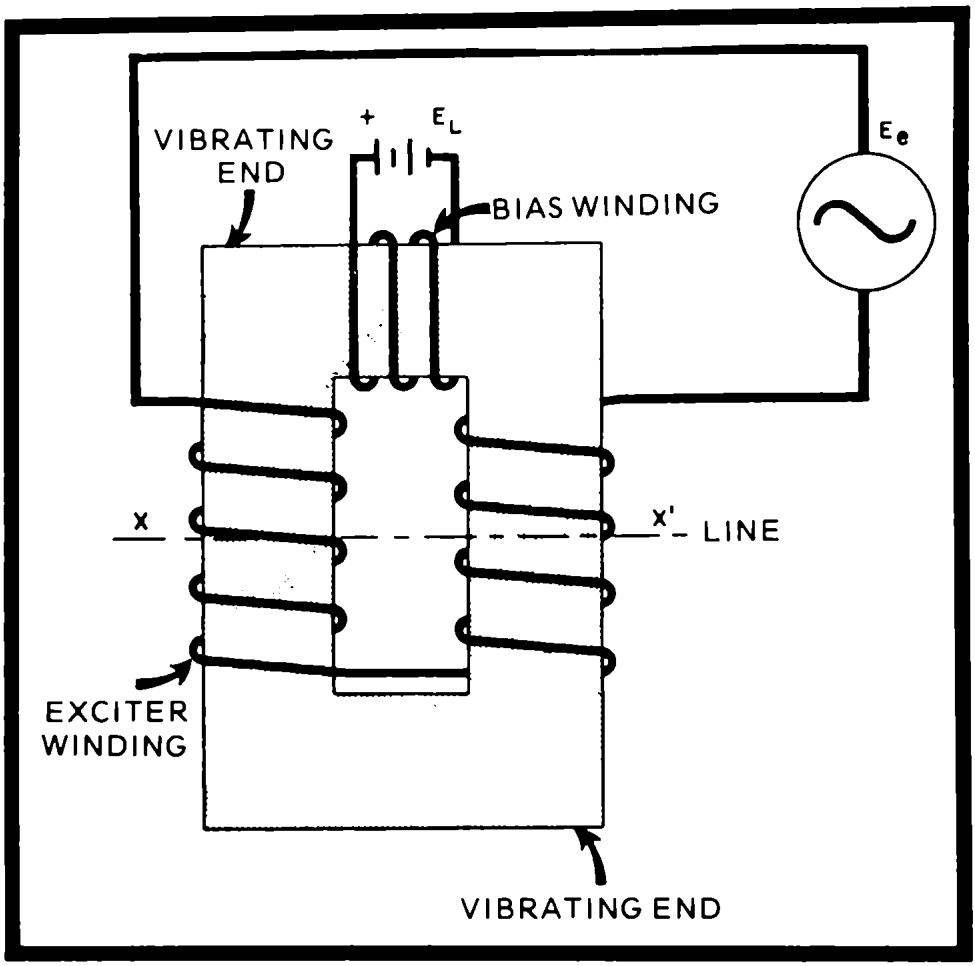


Fig. 6-18. Drawing of the vibrating core in a magnetostrictive ultrasonic generator.

a fair number of other reactions, among others in colloid chemistry, are accelerated if not made possible by ultrasonic treatment.

Intense ultrasonic agitation makes for an intimate mixture of powders and liquid solutions, useful in preparing photographic emulsions, paints and some pharmaceutical products. By ultrasonic agitation, it has been possible to prepare a colloidal emulsion of fine mercury droplets in water—two substances that do not mix normally.

Subjecting a degreasing or cleaning bath for delicate mechanical pieces to ultrasonic vibrations accelerates and intensifies the cleaning action, since grease and other foreign particles are quickly torn away. Experimental ultrasonic cloth washers have yielded encouraging results. There are a good many applications in the food and dairy industries: sterilization of water, milk and soft drinks by killing bacteria, accelerated aging of wine, and controlled fermentation and maturation of some cheeses.

ULTRASONIC MACHINING

Ultrasonic tools are especially attractive for cutting or machining holes of intricate shapes and cavities into very hard and brittle materials (tungsten carbide, germanium and silicon, ceramics, etc.). An ultrasonic drill (Fig. 6-19) uses a magnetostrictive transducer as described above, firmly clamped at its nodel line xx' and a metallic coupling cone of logarithmic shape bonded to the lower end of the transducer body. The coupling cone holds the cutting tip. The workpiece is immersed in a suspension of an abrasive powder in water, and the cutting tip vibrates very near the work without actually touching it. As the tip vibrates up and down, the fluid transmits its movements to the abrasive particles that gradually sink a hole in the work, corresponding exactly to the shape of the tip. These particles thus do the cutting job, but the liquid is necessary to transmit and vibratory energy and to cool the work and the tip. Because the cutting operation is performed by the abrasive rather than the tip, the latter can be made of a comparatively soft and easily machinable material such as copper. Since hard materials are more rapidly abraded than soft ones, the tool wears much less than the work.

Ultrasonic drills, generally automatic, are fitted with a precision tool feed that lowers the tool into the eroded work while continually optimizing the gap between tool and work. This is accomplished by a special servomechanism. Ultrasonic drills usually deliver several hundred watts of vibratory power at frequencies ranging from 15 to 50 kHz.

ULTRASONIC SOLDERING AND WELDING

Aluminum and its alloys are rather impossible to solder by conventional methods, a servere drawback for electronic chassis wiring. The oxide film that immediately covers a bare aluminum surface and prevents soldering, however, can be broken up by the vibrating tip of an ultrasonic soldering iron which produces many tiny vapor bubbles that agitate the molten solder. Since the solder-covered aluminum surface cannot oxidize again, soldering now becomes possible.

The mechanism of an ultrasonic soldering iron and the diagram of the driving generator are shown in Fig. 6-20. The transducer is seated on the handle in its nodal plane xx', so that the tip vibrates in regard to the handle or supporting fixture. The soldering tip, heated as usual, is mounted to the small end of the exponential transition cone, whose base is bonded to the free end of the magnetostrictive block. Winding L1 is fed the AC exciting power, and a DC bias voltage is ap-

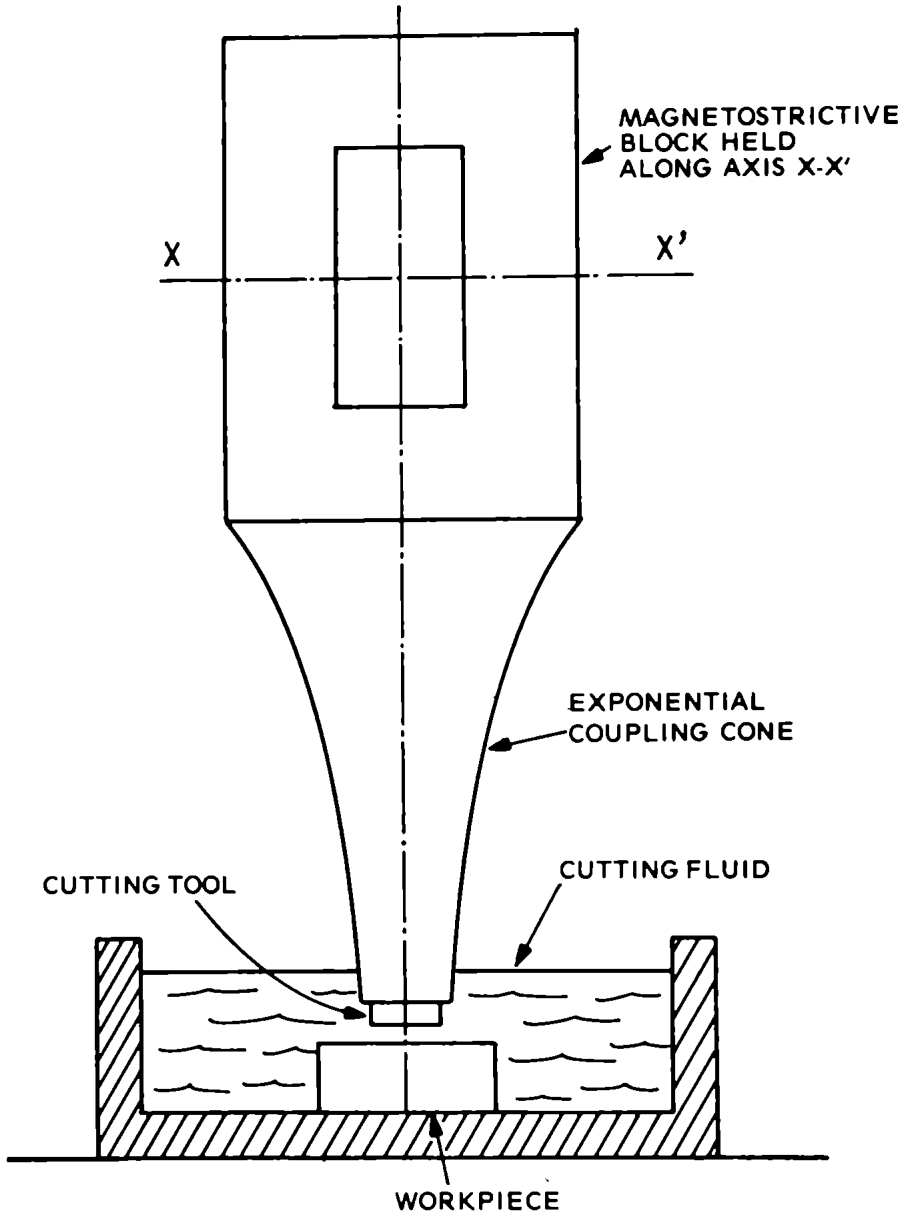


Fig. 6-19. Sketch of a magnetostrictive drill or cutting tool.

plied to winding L2. The other end of the magnetostrictive stack faces a small C-core comprising a permanent magnet, laminated poles and a feedback winding (L3).

The transducer is energized by a power amplifier driving L1. The vibrations of the free end of the transducer induce a voltage of the same frequency in pickup coil L3, and this voltage is injected into transistor Q1's base to build up and sustain oscillations. By this means, the electrical oscillation

frequency automatically tunes to the mechanical resonance frequency for highest efficiency. Capacitor C tuned L3 to the operating frequency and eliminates harmonics.

Ultrasonic welding is a recent technique using the mechanical vibratory energy of a cold spherical welding tip to bond similar and also dissimilar metals. While the mechanism of this process is not completely understood as yet, it seems that a solid-state fusion resulting from plastic deformation, or cold welding, takes place. The equipment used is basically similar to the soldering iron discussed above, but the tip, of

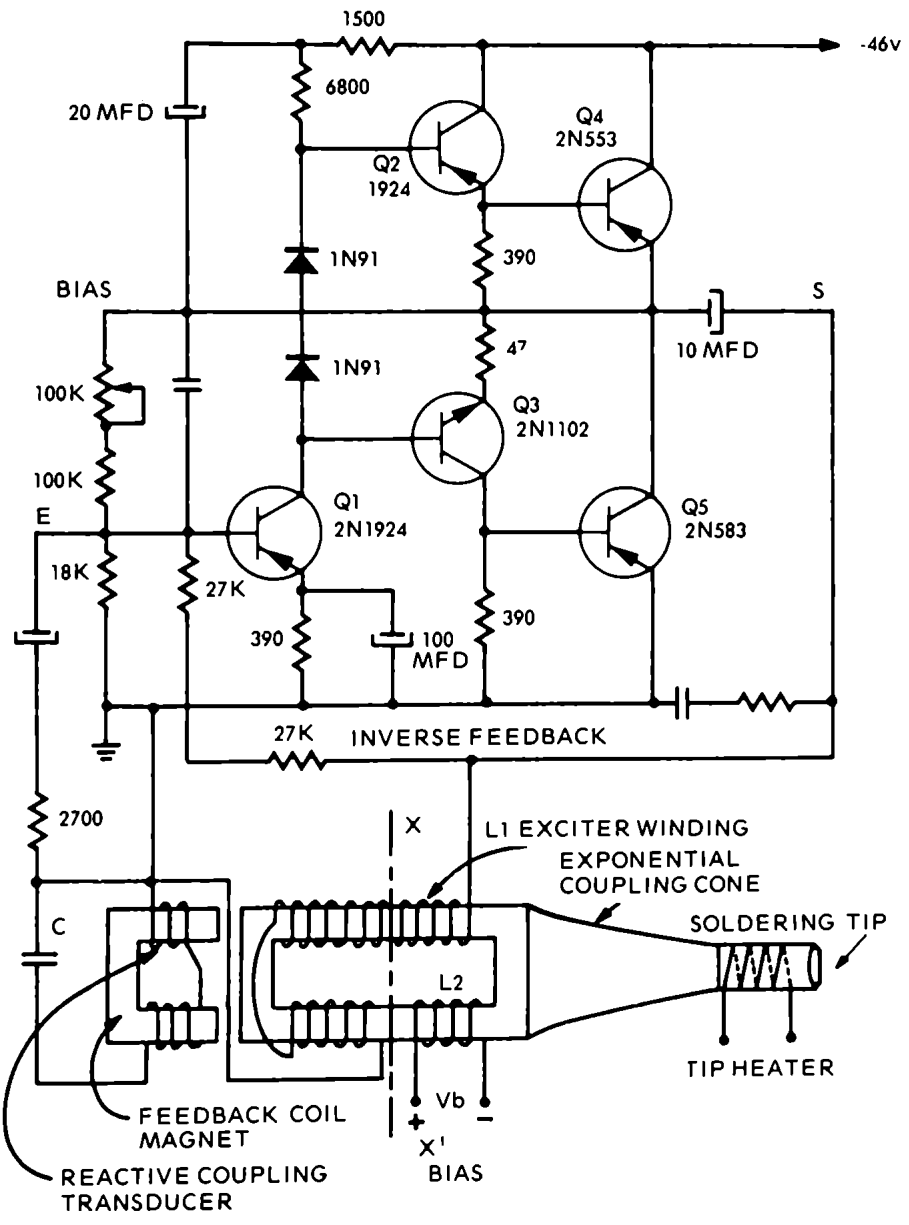


Fig. 6-20. Schematic of an ultrasonic soldering iron.

course, needs no heater. Small hand tools provide some 25 watts of vibratory power and bench-type welders from 300 to 4,000 watts. The operating frequency usually ranges from 15 to 50 kHz.

SPARK MACHINING

Just like ultrasonic machining, electric spark machining is an interesting method for cutting openings of intricate or irregular shape into very hard metals such as carbides (insulators cannot be machined by this process). Here, too, the cutting tip can be soft, since it actually does not touch the workpiece, and the entire job is immersed in a liquid; but there the similarity ends, for the process is entirely different.

In the spark machining process, some 40 to 400 volts is applied between the tool and workpiece, which are separated by a very narrow gap (about .001 inch) filled with a dielectric fluid such as transformer oil. In the first spark machines, a capacitor was charged by a DC voltage through a resistor to build up an increasing voltage until the dielectric breaks down and a spark discharge takes place. Since the work is connected to the positive pole (anode), tiny particles are torn off at the spark impact, driven towards the (negative) tool and flushed away by the fluid. This gradually erodes the workpiece to conform exactly with the shape of the tool. Since the spark always flashes between the points separated by the smallest gap length, the particles are removed very uniformly.

The dielectric fluid serves the triple role of (1) washing away the torn-off particles, (2) cooling both the work and the tool, and (3) quenching the spark to avoid arcing that that could damage the worked surface. The applied voltage retains the same polarity to insure work rather than tool erosion. As soft metals are less eroded than hard ones, the tool bit usually is made of copper or soft steel, materials easy to machine.

Besides the cutting process that is purely electronic, electronic techniques are used in at least two parts of a spark machine: for spark voltage generation and for automatic tool feed to control the gap length. Furthermore, there may be a programmed work-positioning system as used in other machine tools.) Since particle removal occurs only during the spark action, efficiency is greatly increased by increasing the sparking frequency. Machines now feature spark repetition frequencies of up to 50 kHz, and special models work at still higher frequencies. Spark discharge currents range from about 100 to 20,000 A, but because of their short duration, the average power absorbed is moderate.

A typical medium-frequency spark-voltage generator is schematically represented in Fig. 6-21. The supply transformer voltage, rectified by V1, charges capacitor C through resistor R. C supplies anode voltage to thyratron V2, whose cathode circuit contains the primary of pulse transformer T2. The T2 secondary is connected to the work and tool system. V2, biased to be normally nonconductive, is fired by trigger pulses applied to its grid. The fired thyratron applies the energy stored in C to the pulse transformer primary, starting an oscillating discharge. Damper diode V3 kills the negative half-cycle of this damped oscillation to preserve one-polarity output and reduce tool wear. Many other spark-generating devices are in use, including the rotating-wheel spark gap. Radar-type delay line charging and discharging systems are also used.

Just as in ultrasonic machining, there is a micrometric tool feed that maintains constant the very small gap length as the tool eats into the work. The servomechanism operating this micrometric feed obviously has to be very accurate and backlash-free. Since the spark voltage or current is a function of gap length, an electric signal for feed control is available.

An interesting special application of spark machining is cutting threads into tungsten carbide or other very hard stock, for customary taps wear down more than the work. The spark machine gradually lowers a slowly rotating screw made of soft material into the work, so that the particles facing the screw threads are removed by the spark process. This

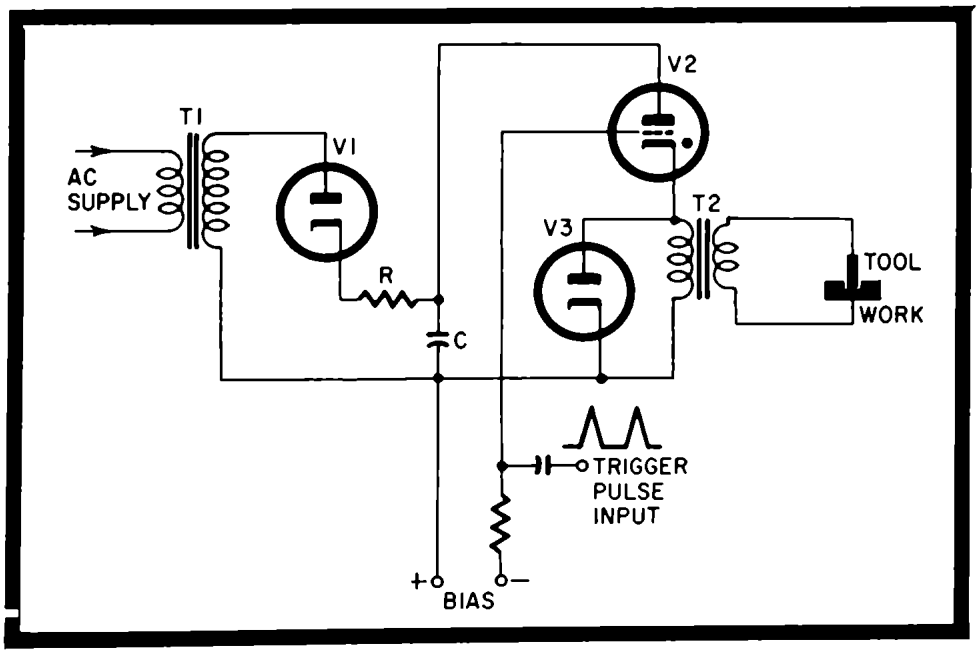


Fig. 6-21. Circuit of a spark-machining pulse generator.

requires an accurate coordination of the descending and rotating motions, so that the feed per turn exactly equals the pitch of the screw thread.

EVAPORATION, DRILLING AND WELDING BY LASER BEAMS

The invention of the laser dates back to only 1960, but its enormous application potential is just beginning to materialize as industrial hardware. Strictly speaking, the laser (Light Amplification by Stimulated Emission of Radiation) is not an electronic device. Its emission is due to a molecular phenomenon, and the emitted beam is composed neither of particles such as electrons nor of vibrating energy, but of mere light. (In fact, light is nothing else than electromagnetic radiation of tremendously high frequency, some 10^{14} Hz, with a wavelength of about 1 micrometer.) The light emitted by a laser is of a special sort: it is **monochromatic** (only one frequency or wavelength is present) and **coherent** (all beam components are in-phase). Because of these two properties, a laser beam can be focussed in an extremely small spot where its energetic density can be amazingly high. The same beam will not spread, either, when transmitted over great distances. The energy concentration is serviceable for **microprocessing**, while the highly parallel beam is of interest in applications such as communications, optical radar and the like which we will not examine here.

To understand the laser operating principle, consider a material at rest. Most of the atoms it is composed of are at the lowest energetic state called the "fundamental level." An impinging light radiation raises some atoms to a higher energetic state and radiation absorption takes place.

If by an external means (a **pump**), the atom distribution on the energy level diagram of the material can be inverted (this is called a "population inversion"), the radiation absorption phenomenon mentioned above is inversed, too, and there is **radiation amplification**. This is called **stimulated emission**: an impinging photon causes level de-excitation by emitting a radiation. Just like there are absorption lines, there will be **amplification lines**, and its linewidth will be reduced to the extent that a photon emitted by another both have the same frequency and phase.

Building a laser emitter, therefore, it is necessary to raise a great number of atoms from a given level to a high energy level. This is achieved by enclosing the rod-shaped active material in an optical tuned cavity with mirrored end faces so that an initial photon emitted in an axial direction is reflected

back-and-forth a great number of times, thereby producing amplification (Fig. 6-22). One of these mirrors is partly transparent to release a coherent radiation.

The helical tube surrounding the active rod actually is an electronic gas discharge device similar to the flash tube used by photographers. This is the pump, so named because it seems to "pump" the atoms to a higher level; pumping is done by an intense illumination when the tube fires. This leads to pulsed operation, but CW lasers are available, too, they use various gases. Solid-state lasers mostly used for processing comprise active rods made of materials such as YAG (neodymium-doped yttrium-aluminum garnet), ruby and neodymium-doped glass.

The radiation emitted is more than just light. Consider a pulsed laser with one joule of energy whose beam is focussed in a spot 10 micrometers in diameter; that is about one millionth of a square centimeter. Hence, there is an energy density of one million joules per square centimeter. If the pulse has a duration of 1 millisecond, there will be a power density of **one million kilowatts per square centimeter!** It is conceivable that such tremendous energetic densities have considerable thermal effects, although it is "just light."

A laser will weld or machine according to pulse duration. For a given energy and spot surface, the shortest pulses yield the highest power, hence the greatest temperature rise, leading to material volatilization. This is the case of laser beam machining. Longer pulses are good for welding. Table 6-1 summarizes the characteristics of the most used solid-state laser materials.

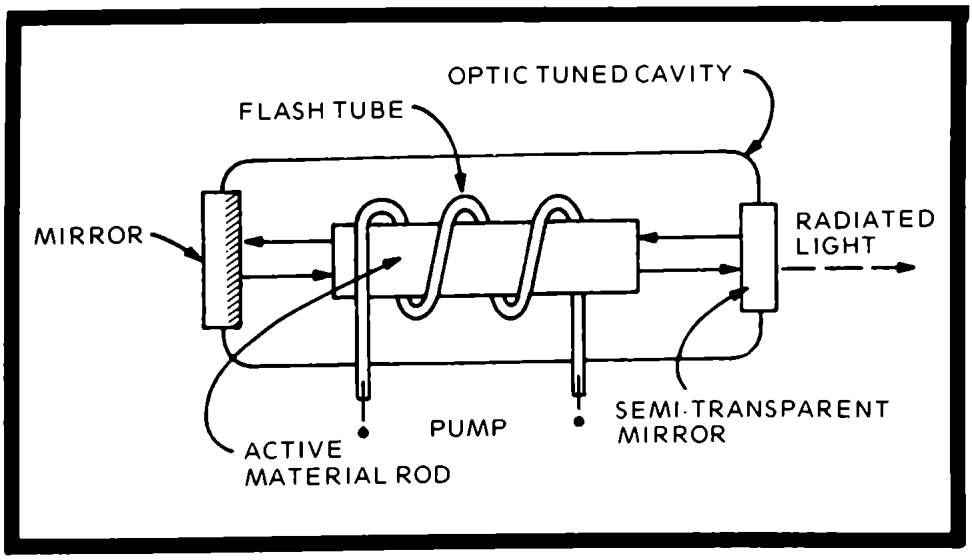


Fig. 6-22. Drawing showing the principle of a solid-state pulsed laser.

Material	Pulse Duration	Energy Per Pulse	Pulses Per Min Application	Application
YAG	50 ns	1 millijoule	48,000	1, 4
	500 us	1 joule	6,000	2
	some ms	some joules	60	3
Ruby	500 us	500 millijoule	6	2, 4
Nd-doped glass	0.5 - 5 ms	1 - 30 joules	10	3

Key to applications: 1. Evaporation; 2, micro-drilling; 3, micro-welding; 4, adjusting resistors and capacitors.

Table 6-1. Characteristics of the most used solid-state laser materials.

One of the first experiments in machining with a laser beam was burning tiny holes in razor blades. Spectacular as it was, manufacturers wondered if there ever would be truly industrial applications justifying commercial laser production. Today, nobody will question any more the promise of laser uses in industry.

Pulsed lasers are specially attractive for pin-point work where the thermic and physico-chemical disturbances in the vicinity of the impact are negligible. As an example of drilling fine holes in almost any material, one pulse laser is used for drilling 80-micrometer holes in watch rubies 1 mm thick, and wire-drawing tungsten carbide dies are conveniently pierced and restored by laser beams.

Evaporating material involves the same process as burning holes, but the beam action is limited to a shallow surface region. By selectively evaporating the conductive layer of resistors and capacitors without damaging the underlying insulating substratum, these components can be accurately adjusted to specified values; this operation can be carried out automatically by bridge control. Similarly, unwanted parts and connections on integrated circuits can be burned away.

Laser welding allows for processing extremely small pieces that can be butt-welded. Many electronic and microelectronic components are eligible for laser spot welding: galvanometer moving coils, relay parts, transistor casings and printed circuits, etc. Fine wires and magnet wire can be directly welded without prior cleaning. Fine spirals can be welded, too, with only a microscopic region being heated.

CHAPTER 7

Electronic Saftey Devices

Besides performing, commanding and supervising production processes, electronic equipment plays an important role in providing full operator safety and in protecting valuable machines and plant. In an automated system, means are usually provided to prevent the machine from damaging or destroying itself, for instance, by overrunning travel stops. If a human operator is included in the system loop, additional equipment is necessary to supervise him and to obviate his wrong decisions or failure to act.

Devising such a fully effective protection system may be a tough psychological problem, for workmen often resent the tutelage of a control that keeps them from doing things their own way, slowing down their work as they often believe. And since the human being is capable of a lot of craftiness, it is difficult to anticipate all his possible inventions to fool the protecting device and to make it inoperative.

PRESS AND SHEAR SAFETY DEVICE

Danger is particularly prevalent with hand-fed power punch presses and shears. If only one button has to be pressed to actuate the machine, one hand will be left free—and endangered. There are some 2-button machines that take both hands to actuate, but this method causes considerable loss of time and the work process may actually need a hand for guiding and maintaining the workpiece. In some cases, operators had had their hands chained so they could not reach into the work space, but, needless to say, people hate being chained like dogs.

A photoelectric safety device may be the best solution to this problem. To protect an area rather than just a line, the light beam issued by a light source is reflected several times by mirrors (M1 and M2) before reaching the photoelectric relay (Fig. 7-1), thus stretching a protective "curtain" between the operator and the danger area. Interrupting the light beam will immediately shut off the machine, and despite continuous operation, there is no risk for the feeding operator. This device does not appear to be easy to fool. (Photoelectric relays are discussed in detail in Chapter 2.)

In the circuit in Fig. 7-2, a low-voltage AC-fed photothyristor, PTH, energizes relay REL. After one or several reflections, the light beam of lamp L strikes the window on the junction of PTH, which fires and energizes REL. Contacts K1 and K2 close, enabling the press to operate. As soon as something, and especially the operator's hand, intercepts the light beam, PTH de-energizes at the end of the half-cycle and will not restrike at the next conducting half-cycle. Contacts K1 and K2 open and the press immediately comes to a standstill.

MAKING SAFETY DEVICES SAFE

Just like any other equipment, a safety device can possibly fail, and the protective system must be designed to

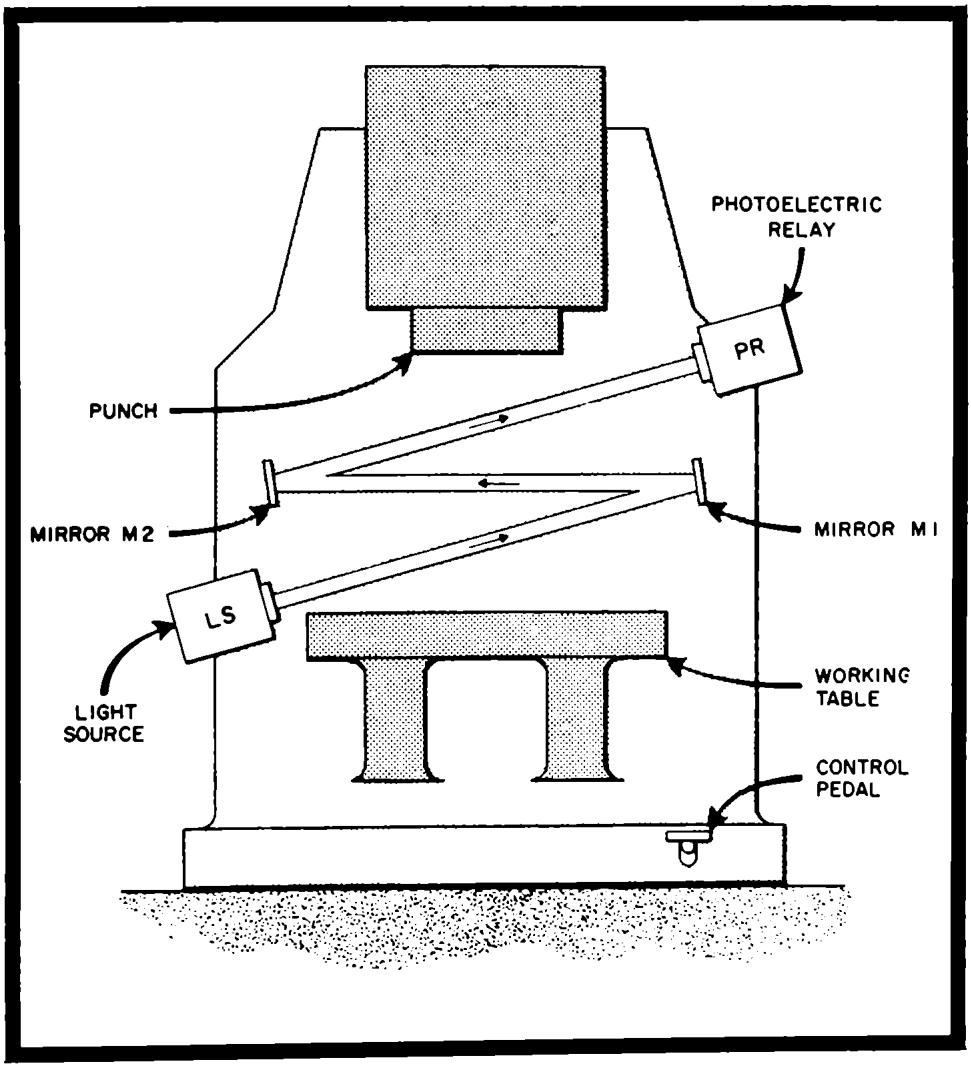


Fig. 7-1. The operator of a punch press can be protected by a light-beam "curtain."

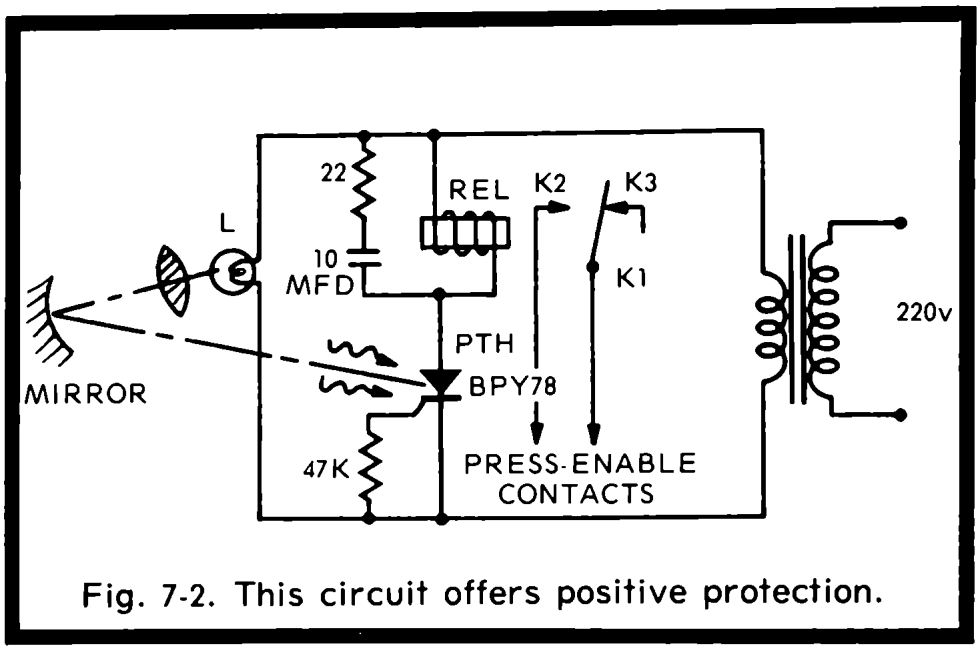


Fig. 7-2. This circuit offers positive protection.

avoid catastrophic consequences should such a failure occur. Basically, the system must provide **positive protection**. This means that any conceivable failure will stop the machine rather than allow uncontrolled operation. This fundamental safety principle has always been applied to railway signals that automatically turn to red should something fail; a delay is preferred to the risk of a collision or derailing. The safety device in Fig. 7-2 conforms with this positive protection principle. To enable the machine to operate, the relay must attract its armature and close contacts K1 and K2. In case of current failure (due, for instance, to a blown fuse), dirty contacts, or a burned-out lamp, the machine is inoperative and no accident is possible.

But just how safe are safety devices? Consider the very simple photoelectric relay in Fig. 7-3, comprising a photoconductor (P) series-connected with relay coil REL (A) or in parallel with the coil (B). In the series circuit (B), illumination lowers the photoconductor's resistance so that circuit current becomes strong enough to energize the relay, closing make contacts K1 and K2. These contacts are used to close the "enable" contacts of the machine control. This is positive protection, expressed by: **the machine can operate only if the photoelectric device is illuminated.**

Now consider the parallel circuit in Fig. 7-3B. REL is normally energized through (low-value) series resistor R, but illumination sufficiently reduces photoconductor resistance to virtually short-circuit the relay coil, closing break contacts K1 and K3. Hence, contacts K1 and K3 can be used to close the "enable" circuit of the machine control, for they close if

nothing intercepts the light ray. However, this is negative protection, for K1 and K3 are the rest position of the non-excited relay and they will be closed, too, in case of a current failure. Thus. illumination of the photocell is not an exclusive operating condition of the machine, and the operator will be induced to believe he is safe when actually he is not.

A machine must be permitted to operate only if all safety conditions are fulfilled simultaneously; this is expressed by

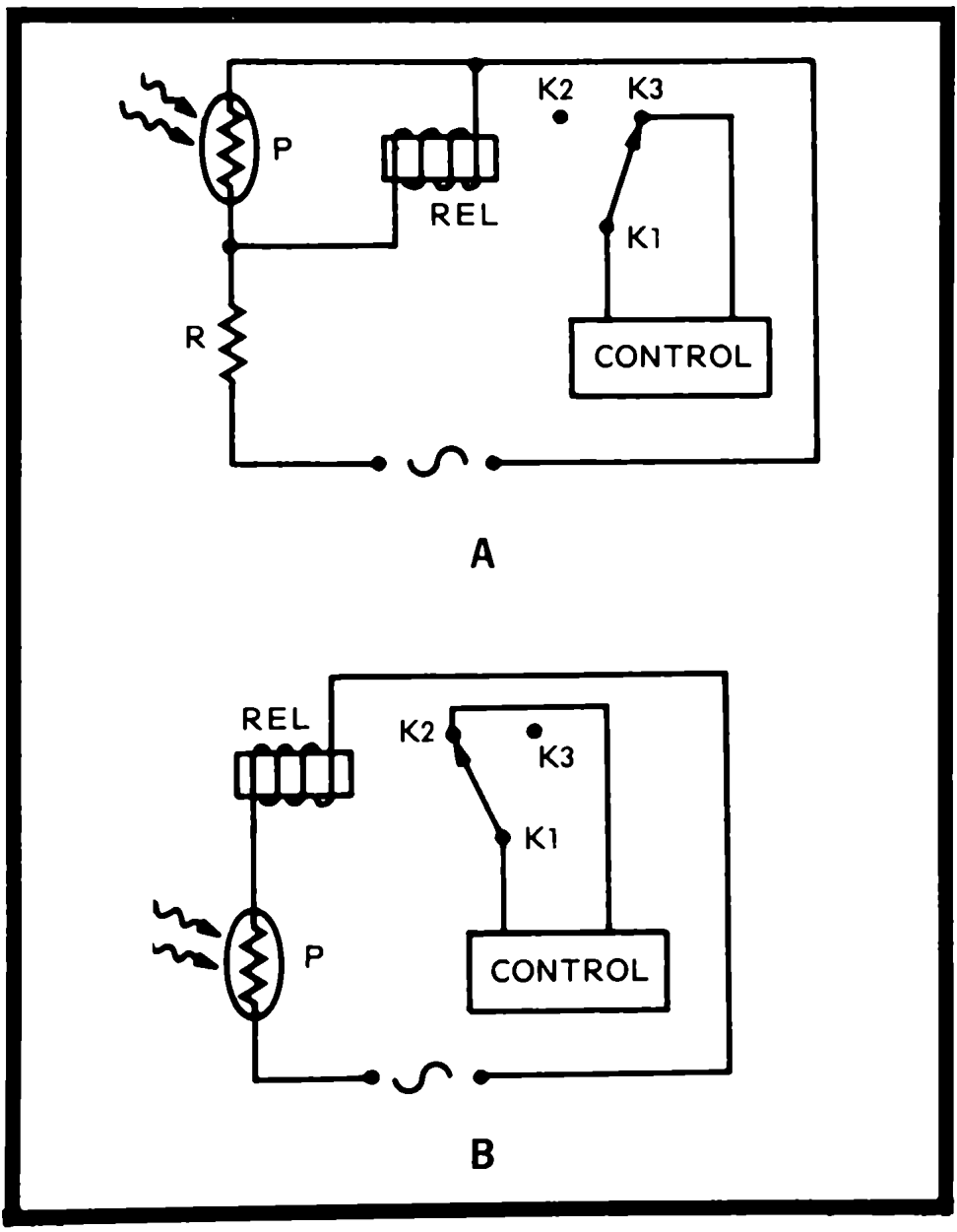


Fig. 7-3. The circuit at the top provides positive protection because its associated device cannot operate unless this circuit is functioning normally. The circuit at the bottom offers negative protection.

the AND function in logic algebra and carried down to earth by closing series-connected switches. In a microwave oven, for instance, the magnetron heater is on in the stand-by position, so that heat immediately is produced upon connecting the B+ supply. This connection, however, is possible only if the oven door is closed (for operator safety) AND the cooling system is on (fan or circulating water) AND the tube is pre-heated AND the foot-pedal or button is pressed to order heat production. This set of conditions to be fulfilled is by no means limitative. This requirement also explains why an industrial electronics diagram, though basically simple, may appear inextricable due to the presence of a number of safety circuits that do not alter the operating principle of the machine. A serviceman called to repair an electronic machine breakdown will gain time if he bears in mind all these series-connected safety switches, since a defective contact on one of these switches is enough to forbid machine operation.

FLAME-FAILURE PROTECTION

In industrial furnaces, a protection system is required to stop fuel arrival, if the flame goes out or if the gas or fuel does not ignite, to avoid fuel flowing into the hot hearth where it could cause an explosion. This control has to interrupt the circuit of the solenoid gas or fuel-supply valve or the pump motor immediately when the flame goes out or fails to appear.

A photoelectric device trained on the hearth can be used as a flame detector. Because of the radiated heat, an optical filter must be interposed between hearth and photoelectric detector to prevent it from being responsive to the mere heat radiation of the hot hearth after flame extinction, obviously a catastrophic error.

The flame detector in Fig. 7-4 uses a photoconductor PR series-connected with the coil of relay REL. While a photoconductor works equally well on DC and on AC, DC is better for operating relays (if they are not of a special AC-type). To grant quiet chatterless relay operation, the AC supply voltage is rectified by diode D and a smoothing capacitor is shunted across the relay coil.

There are flame extinction detectors using photoelectric devices that are responsive to short-wavelength light (from blue to ultraviolet); gas-filled emission-type photocells having this type of spectral sensitivity are available, too. As such wavelengths exist only in flames, and not at all in radiation from hot bodies, detectors of this sort are, of course, a natural to detect flame extinction by optical means.

Another approach in furnace protection is based on electric flame conduction. A ring-shaped wire made of a

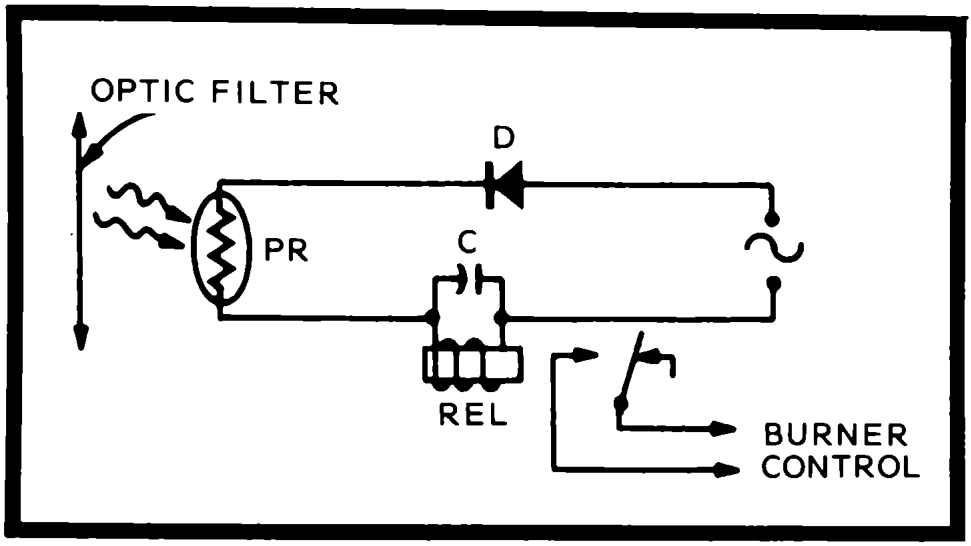


Fig. 7-4. Photoelectric flame detector circuit for furnace protection.

refractory metal, such as tungsten, acts as a "hot" electrode (both meanings) exposed to the flame; the grounded burner forms the other pole. Since flames are high-impedance mediums, the flame probe is connected to a high-impedance input stage using an FET (Fig. 7-5). The FET input electrode is biased by a highly resistive input voltage divider and is grounded by the flame. If the flame goes out, the bias of the source-follower FET-controlled UJT is reduced, thyristor TH is fired, energizing relay REL and opening the burner "enable" contacts. This stops fuel arrival. The circuit, and especially TH, is DC-fed to make sure that the shut-off fuel supply remains so to avoid a possible explosion. A new operating cycle can be started by pressing the reset button to short-circuit TH in order to de-energize it.

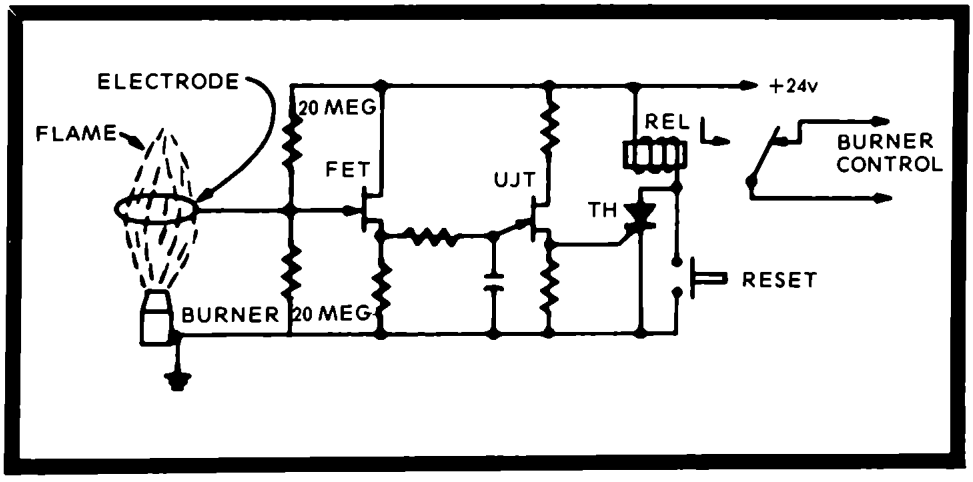


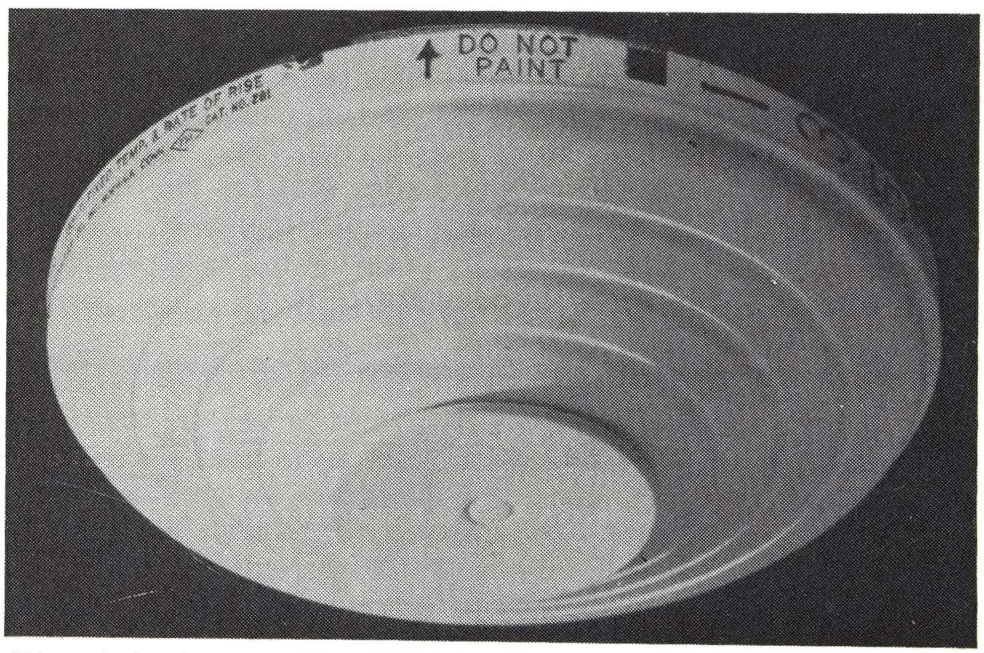
Fig. 7-5. Furnace flame conduction detector circuit.

If the furnace is automatically controlled, everything stops if the flame goes out. Then, after a preset time, the fuel arrives again, and the electric igniter is struck. If there still is no flame, there will be another delay followed by another ignition trial. After three or four unsuccessful trials, the control gives up (only very small spilled fuel quantities can be tolerated in a hearth) and signals the breakdown to the supervisor.

FIRE AND SMOKE DETECTORS

Such instruments are actually temperature rise, infrared radiation or smoke detectors. They usually sound an alarm if there are signs that a fire has broken out and sometimes also are used to trigger a sprinkler system. Such systems are very sensitive: in one case at least, the smoke of a cigarette lighted by an imprudent visitor is said to have triggered a fire-alarm and abundantly irrigated the culprit.

The smoke detector in Fig. 7-6 is simply an application of the differential light-sensitive control in Fig. 2-56. A light box issues two light beams that strike the junctions of phototransistors PT1 and PT2, after traveling a distance through free air. One beam, however, is surrounded by an



Fire detector designed to survey a maximum area of 2500 square feet. It is a combination fixed-temperature, rate-of-rise unit. The rate-of-rise element is self-restoring; activation of the fixed-temperature element causes the center disc to drop to a vertical position. Manufactured by Edwards Co.

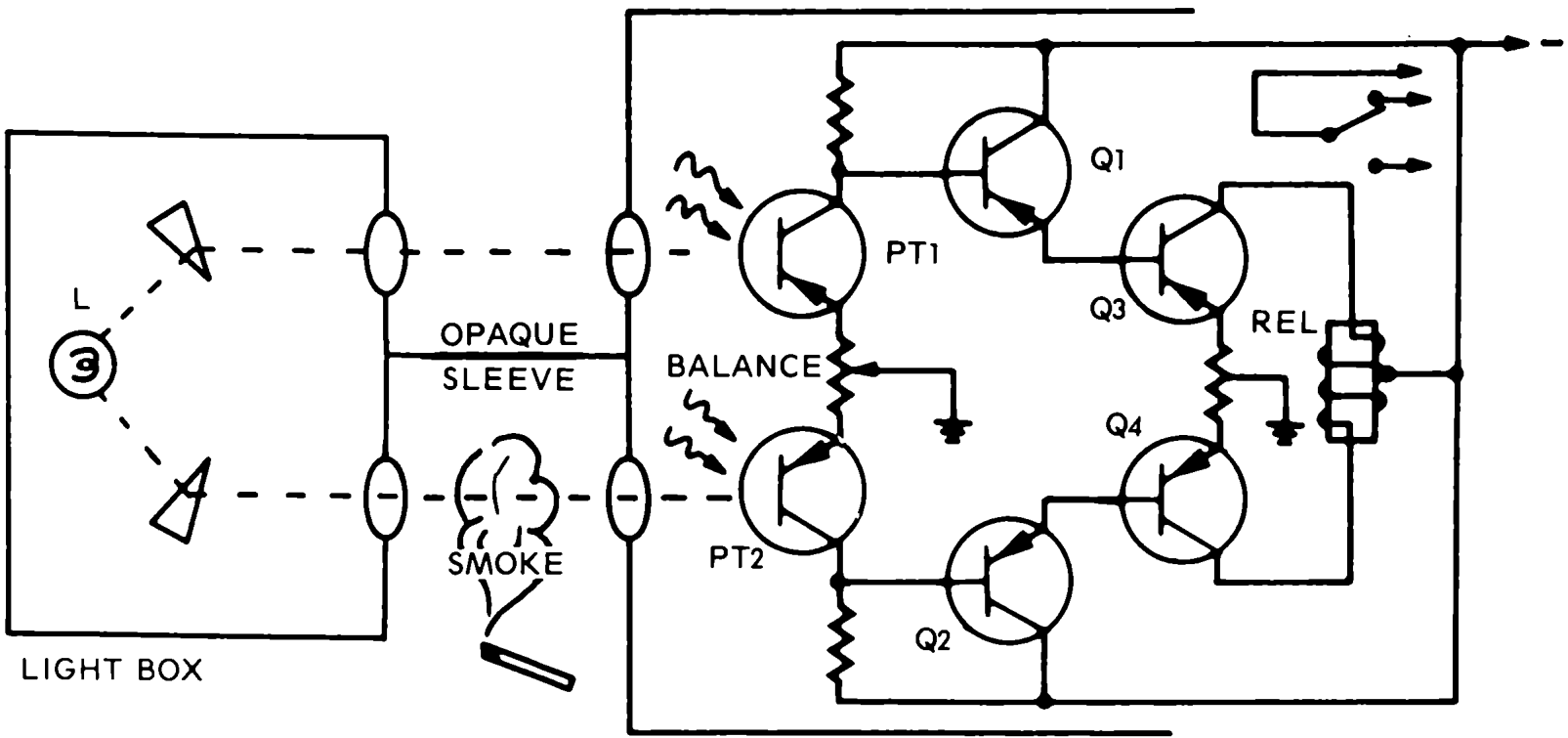


Fig. 7-6. Differential photoelectric smoke detector circuit.

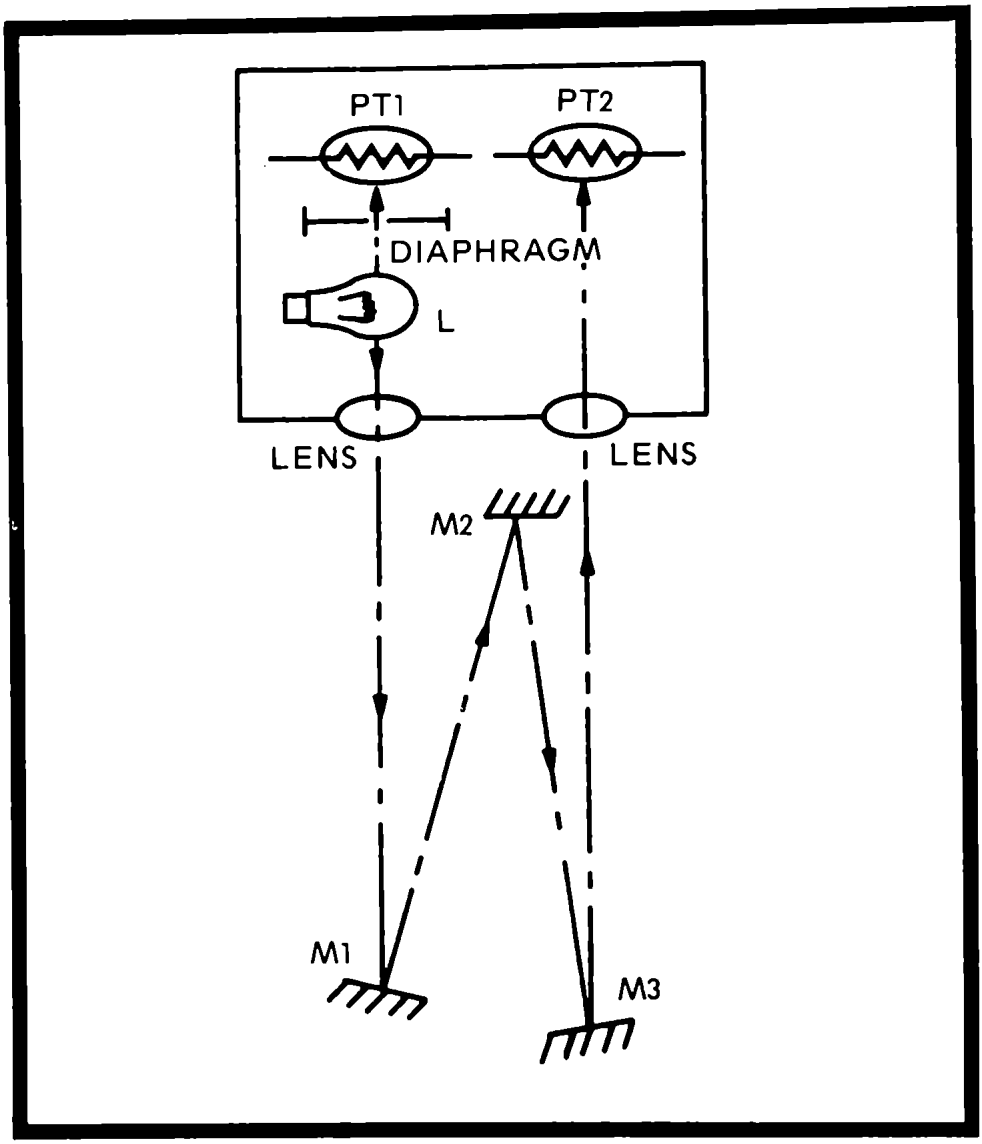


Fig. 7-7. Improved smoke detector circuit for area supervision.

opaque tubing, preventing any interception along its travel. Thus, the presence of smoke or dust clouds in the area overseen will partially obscure but one junction. Due to its symmetry, this device features a high differential illumination sensitivity, while lamp brightness level variations influencing both channels have virtually no effect. The unbalance between the currents is amplified by transistors Q1 through Q4, which energizes center-tapped biased relay REL triggers the fire alarm. The relay sometimes is fitted with a holding contact in order to sustain the alarm, if the cause of its initation vanishes, until manually stopped by the guardian thereby obliged to oversee the area.

The arrangement in Fig. 7-7 dispenses with the opaque tubing. Lamp L is placed within the photoelectric receiver box and directly illuminates PT1 (a gray filter or diaphragm may be needed between L and PT1 to balance phototransistor illumination). A light flux leaves the box on the opposite side and is focussed by a convex lens and reflected back-and-forth by mirrors M1 through M3, allowing for area coverage before illuminating the PT2 junction on. Since only the PT2 beam is exposed to open air, only this phototransistor's current can be altered by smoke particles.

Fire detectors sensitive to infrared radiation use lead sulphide photoconductors or photoemissive tubes responsive to such long-wavelength radiations; a red filter is provided to screen out visible light. Several circuits described in the foregoing are serviceable for this application. Similar instruments are also extensively used in "hot box" detection in railway operations. By amplifying the detector output, it has been possible to detect the flame of a cigarette lighter several miles away, and the heat radiated by a human body at some hundreds yards distance. But to obtain such tremendous

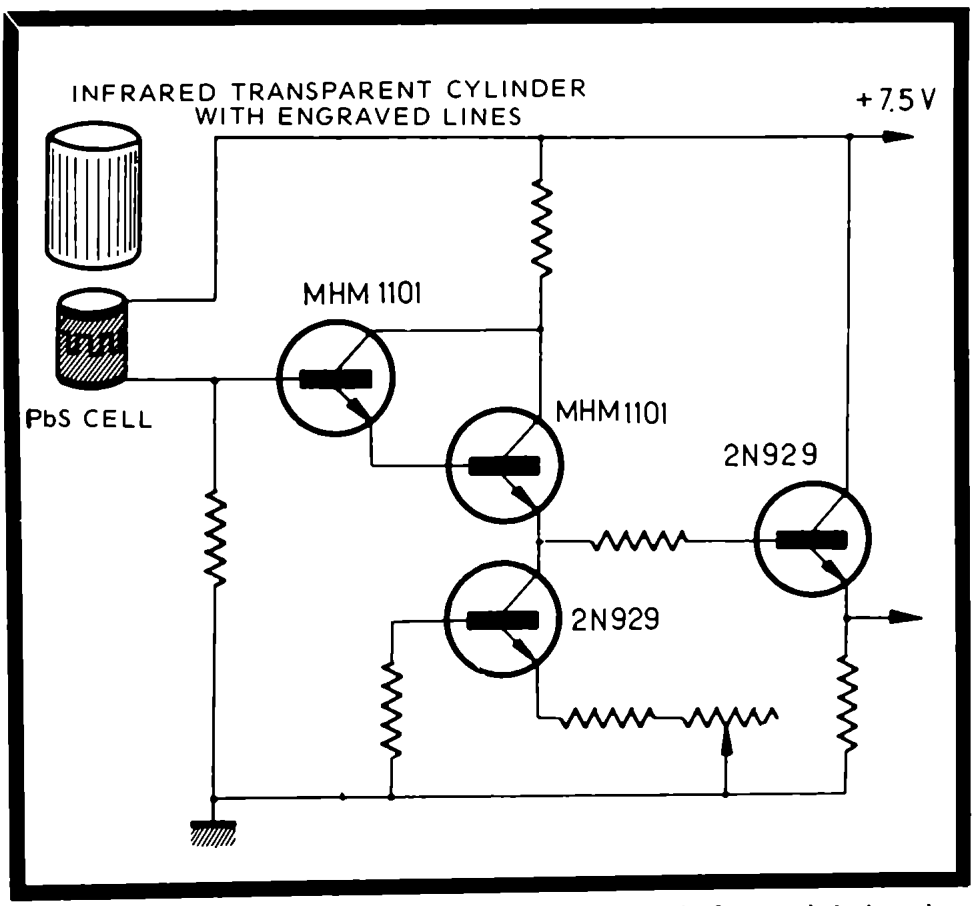
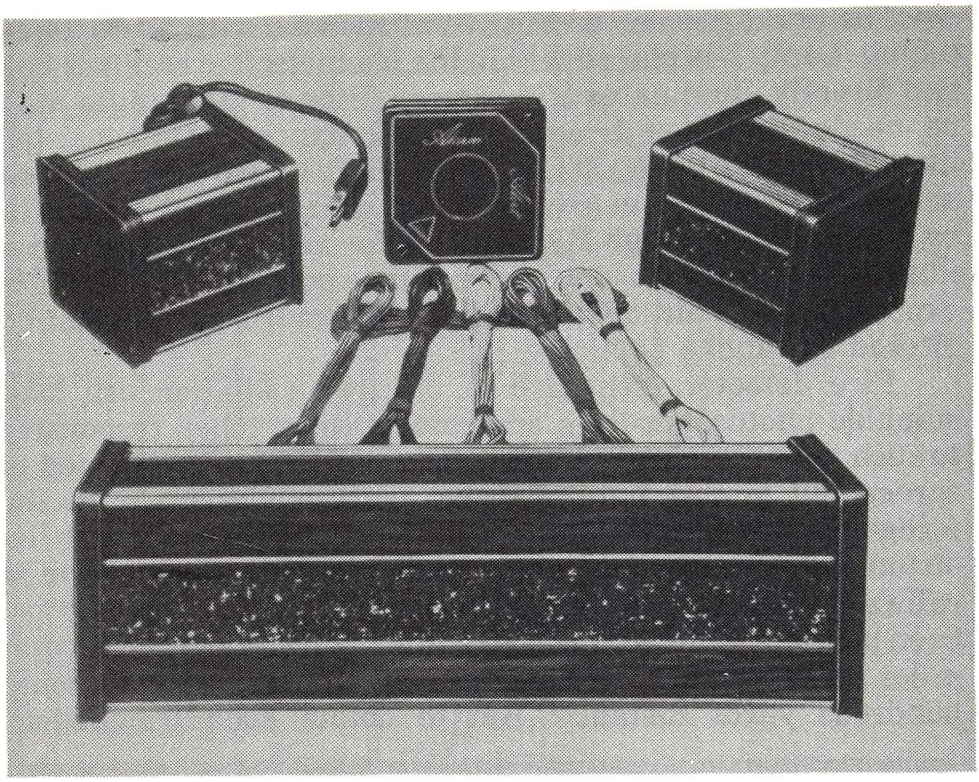


Fig. 7-8. Schematic of a self-chopping infrared intruder alarm.



Intrusion alarm capable of monitoring an area up to 300 square feet. It covers the sensitive area with ultrasonic sound waves. Manufactured by Delta Products Co.

radiation sensitivity, the photoelectric device has to be worked at very low temperatures to sufficiently reduce its thermal noise, so that the very small signal clearly stands out.

INTRUDER ALARM SYSTEMS

Various approaches can be used to detect intruders in a room or in a limited area under surveillance: interception of a light beam (to be concealed, an infrared beam is preferred), listening to sounds and noises, detection of a light even faint in a dark environment, or a capacitive variation due to a human body coming close to some type of antenna. All of these operating principles are serviceable and put to use in actual equipment.

Some recent approach of interest is the detection of the infrared radiation (wavelength 9.3 micrometers) a human being continuously emits (equivalent to about 500w of power). This type of radiation is readily detected by a lead sulphide photoconductor. The resulting signal, however, is small, and the drift of a direct-coupled amplifier is liable to falsely trigger an alarm. This shortcoming, of course, can be over-

come with a chopper amplifier or by chopping the radiation by a mechanical means such as the apertured disc in Fig. 2-57.

A very clever solution to this problem has been proposed (W. E. Osborne, **Electronics**, August 1963, pp. 26-27), using the motion of the intruder to generate this modulation. The technique involves a photoconductive coating on a cylindrical base, surrounded by another cylinder made of a material transparent to the infrared radiation considered. Uniformly spaced vertical lines are engraved in the outer surface of this cylinder, so that it behaves as an optic modulator (Fig. 7-8). With a grid of 25 lines per millimeter, a person moving at a speed of 1 foot per second at a distance of 10 feet from the detector generates a modulation at a frequency of about 6 Hz. If he moves quicker, the frequency, of course, is higher. This signal is amplified by a transistor amplifier with a high-impedance (about 100,000K) Darlington input stage. The feedback-stabilized amplifier controls an alarm relay by means of a thyristor. According to the inventor, this detector has a range of approximately 23 to 27 yards in a most adverse ambient (rain, fog), which considerably attenuates the emitted radiation.

PROXIMITY SWITCH FOR PREVENTING COLLISIONS

When driving large overhead cranes, the operator often does not have an unobstructed view of the scene of action, and means, therefore should be provided to prevent parts of the crane from colliding with stationary masses that may be in the way.

The circuit diagram of an inductive proximity switch designed to stop a crane coming too near to a metallic mass is shown in Fig. 7-9. The device comprises an inductive feedback oscillator with coils L1 and L2 wound on two ferrite rods spaced about 20 cm. Because of the vigorous flux-concentrating effect of ferrites, oscillations will reliably set in even with a gap this wide; but a high-Q circuit is necessary, and any damping by a driving transistor must be avoided. That's why the oscillator uses a MOSFET connected as a tube-type tuned-grid oscillator. To prevent gate insulation breakdown, the gate voltage is limited to 7v by zener diodes ZD1 and ZD2. The source bias circuit, R2-C3, stabilizes the oscillator operation in regard to temperature and supply voltage variations.

When the circuit is oscillating, an AC voltage appears on the drain of MOSFET T1. The negative half cycles control

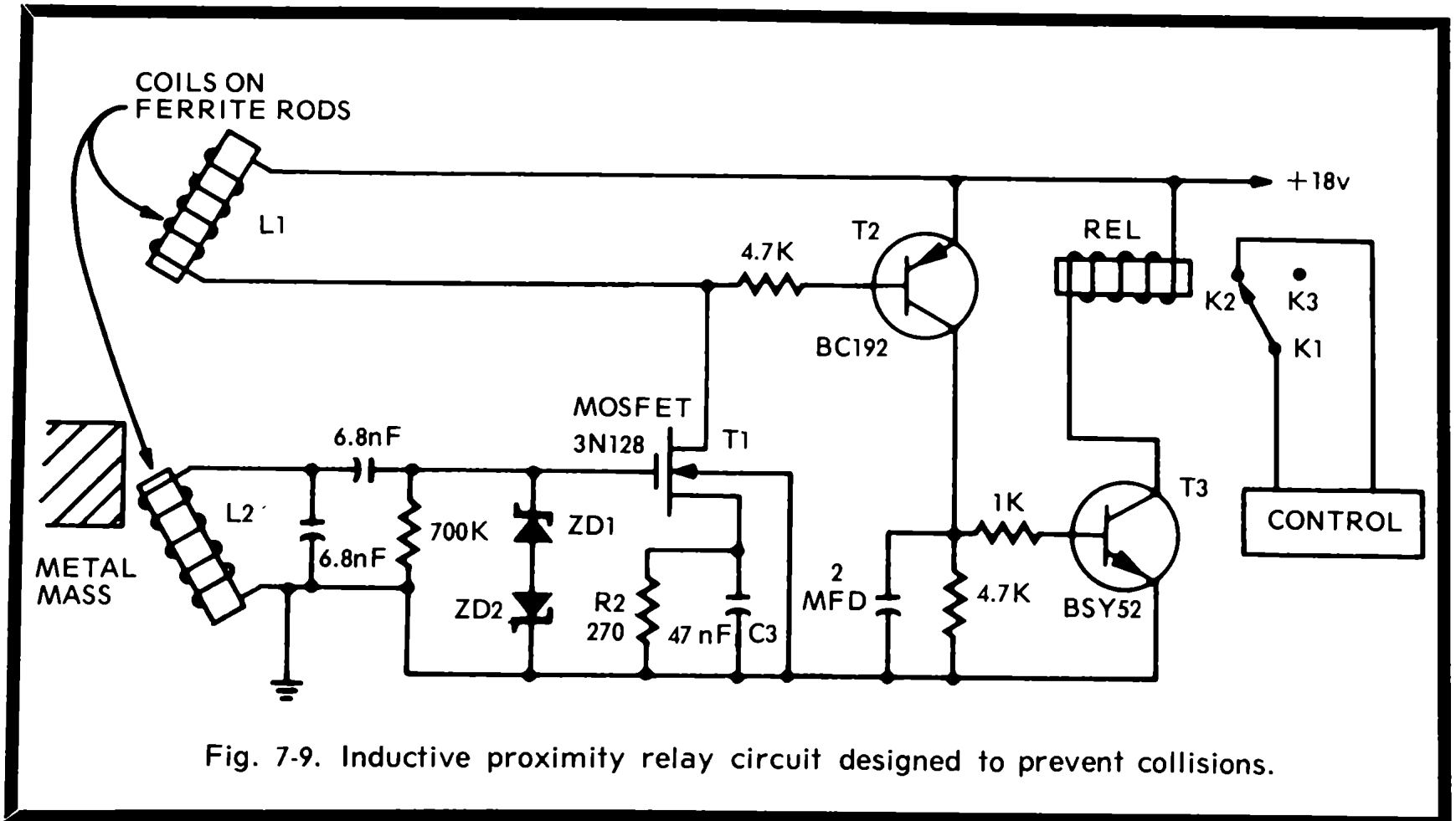


Fig. 7-9. Inductive proximity relay circuit designed to prevent collisions.

transistor T2. This results in positive collector pulses, smoothed to form a DC control voltage which is applied to base of transistor T3. T3 powers relay REL, closing contacts K1 and K2. If a metallic mass comes near the ferrite rod gap, the AC magnetic field is partly consumed by the generation of eddy currents in the nearby metal. This reduces the feedback rate enough to stop the oscillation (the device is adjusted to operate near the threshold of oscillation), REL is de-energized, and contacts K1 and K2 open. At an oscillation frequency of about 120 kHz, the eddy current losses in iron and steel prevail on flux reinforcement. As soon as the metallic body moves away from the proximity detector, oscillation resumes and contacts K1 and K2 close. This is a positive safety device because the "enable" contacts will open should anything fail.

GAMMA AND X-RAY RADIATION MONITORING

Radioactive materials and X-ray machines are used increasingly in industry. The radiations involved are a serious health hazard, especially because human senses are unable to detect them. Radiation detectors, therefore, are extensively used to watch over potentially dangerous areas and to warn operators of excessive radiation levels.

Such radiations are detected by an ionization chamber commonly called a geiger counter. The counter tube is composed of an envelope filled with a controlled gas mixture of definite pressure and coaxial cylinder-shaped outer and inner electrodes (Fig. 7-10). The tube is connected to a high DC voltage supply (about 1,000 volts) through a high series resistance R.

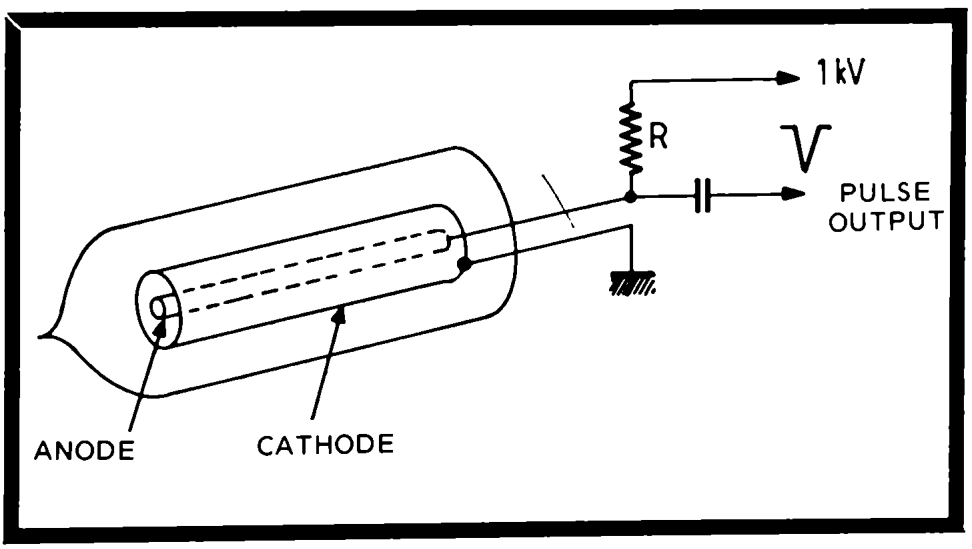


Fig. 7-10. Diagram of a geiger counter tube and its connections.

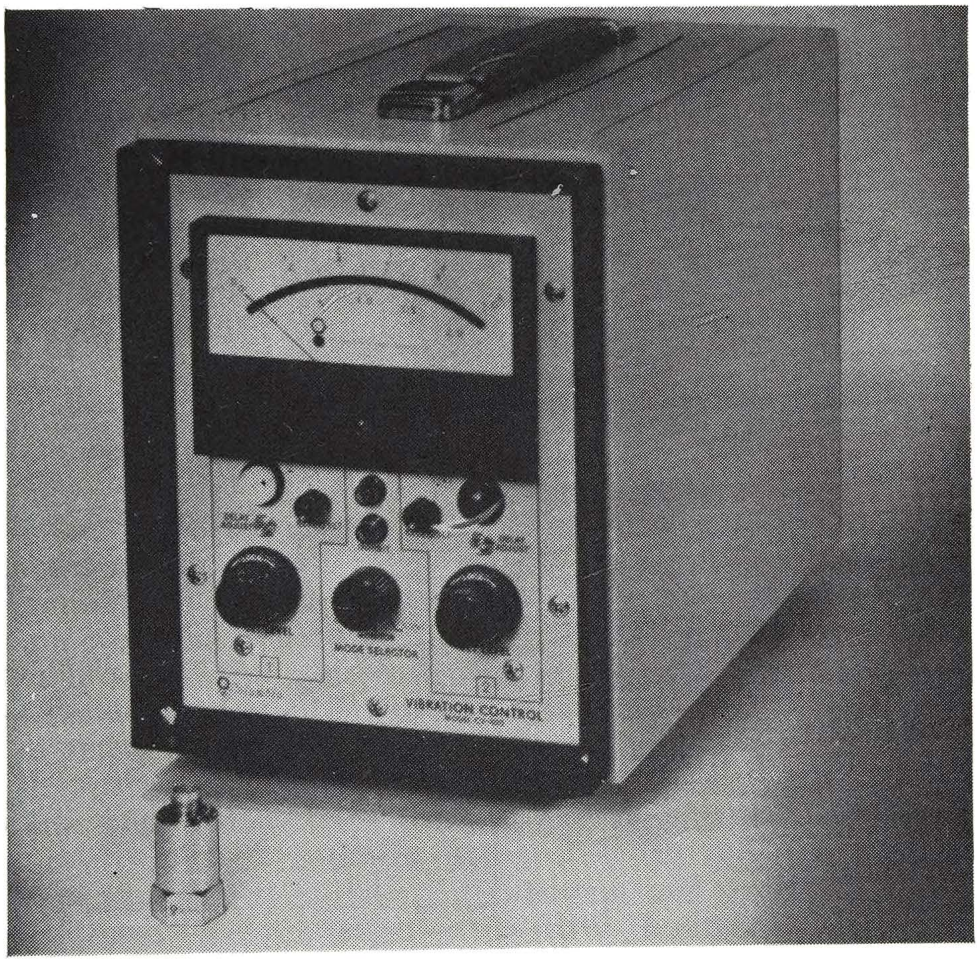
Gaseous oxygen sensor designed to monitor and maintain any desired oxygen level from 2 to 100 percent. Sensor uses two dissimilar electrodes to measure oxygen concentrations. The device is useful in hospitals, laboratories, test chambers, process controls, etc. Made by the Honeywell Apparatus Controls Div.

The tube contains a rare gas such as argon that ionizes when struck by charged particle traversing the envelope; this appreciably lowers the internal resistance of the tube. But if there were only this ionizing gas, the geiger counter would remain ionized as long as voltage was applied, just like any other gas tube. However, this would allow for counting but one particle, the first. A second gas, such as methane, therefore, is introduced in the envelope to **quench** the discharge as quickly as possible, readying the tube for detection of the next particle. Practically, the tube de-ionizes in a fraction of one millisecond. During this interval, it obviously is unable to respond to another particle.

Under these circumstances, each ionizing event results in a pulse across R, and an increasing pulse repetition frequency

indicates an increasing radiation level. You can get a "feel" of this level by listening to the "click" frequency by means of a headset or speaker connected to an amplifier stage (Fig. 7-11); the frequency increases when the detector comes nearer to the source of radioactivity. The pulse amplitude, of course, remains constant.

A visual indication of the radiation level can be obtained by the (analog) pulse counter in Fig. 7-12. The pulses, amplified by transistor T1, charge capacitor C, which discharges between two charging cycles. Both charging and discharging currents pass through the rectifier bridge, so that a DC current proportional to the pulse repetition frequency traverses the meter connected across DC diagonal AB. This circuit is very convenient for portable radiation monitors. The meter scale can be directly calibrated in strokes per second or milliroentgen.



Vibration detector designed to provide automatic measurement, warning, shutdown and protection from excessive bearing and bushing wear. Manufactured by Columbia Research Labs.

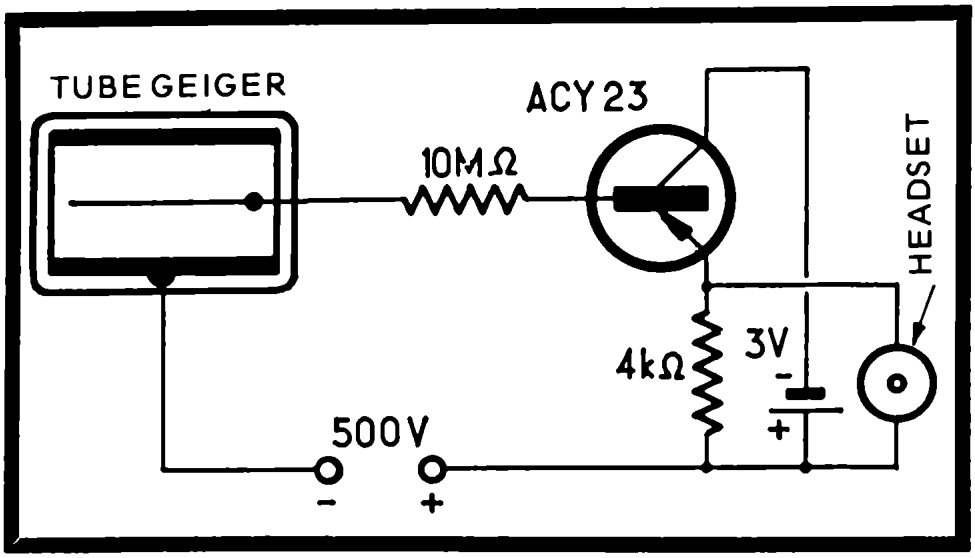


Fig. 7-11. Schematic of a geiger output amplifier.

In radioactive laboratories and production centers, binary and binary-decimal counters of events per unit time (EPUT) are used, mostly to get a view of the radiation level by the flickering frequency of the counter display tubes, especially of the small indicator type. If comparative readings are required, a Nixie-type display is preferred because of the increased reading speed. In such evolved electronic counters, provision is generally made for printing-out radiation counts.

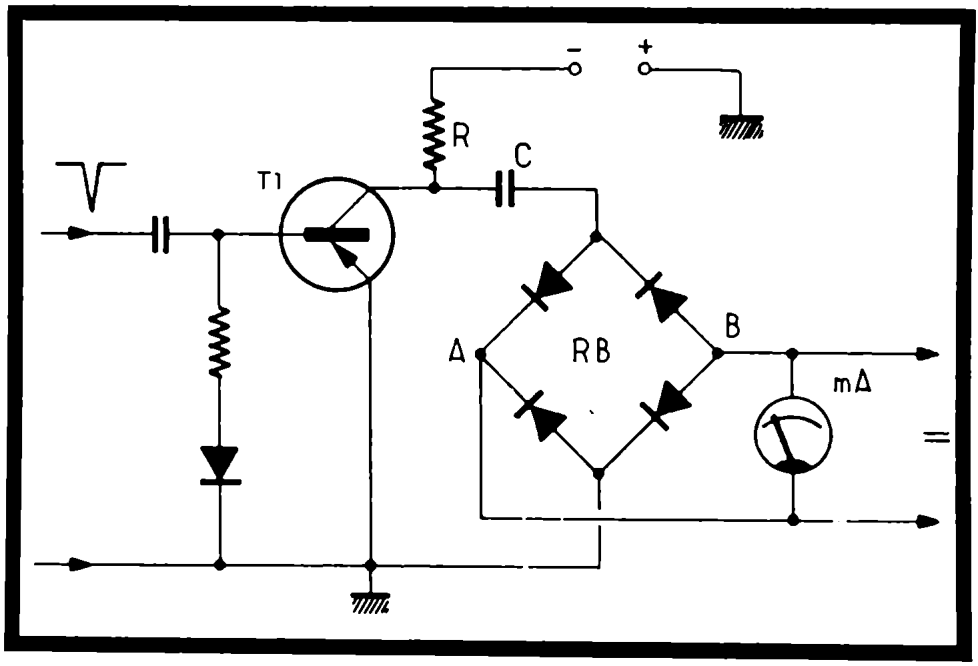


Fig. 7-12. Analog pulse counter circuit for "click" frequency display.

CHAPTER 8

Power Conversion & Control

While the foregoing chapters deal mostly with low-power signal generation and processing, we must now discuss the conversion of electrical energy into energy of another kind (rectifiers, converters and inverters) and into mechanical power (motor control). Power converters or motor controls form the last link of most industrial control systems.

RECTIFIER TYPES AND CONNECTIONS

Rectifiers convert the 60-Hz AC line power, the main source of electrical energy, into DC to operate low-power devices such as electronic circuits as well as larger systems such as DC motors and welders, and electrolytic plants. All of these devices are DC-fed; but powers range from a couple of watts to thousands of kilowatts, with a similar range of voltages and currents. So rectifiers can look quite different, depending on the specific application.

A rectifier transforms the polarity-reversing AC into pulsating DC by passing current only during one half cycle. Besides low-power selenium and copper oxide rectifiers, vacuum and gas-filled diodes were used solely three decades ago. Nowadays, germanium and silicon rectifiers have virtually superseded electron tubes in all but a few special applications. Technologic progress is continuously increasing the ratings of solid-state components. Compact devices with a rectifying ability of thousands of amps, capable of withstanding thousands of volts, are readily available now. Semiconductor rectifiers feature a rather low direct resistance in comparison with vacuum and even gas tubes, resulting in a much lower voltage drop, less heat dissipation, and, thus, better efficiency, especially at low voltages. Furthermore, no separate heater winding is necessary, so no heater power is lost.

A semiconductor rectifier is more compact, more rugged and more reliable than the corresponding tube, and its

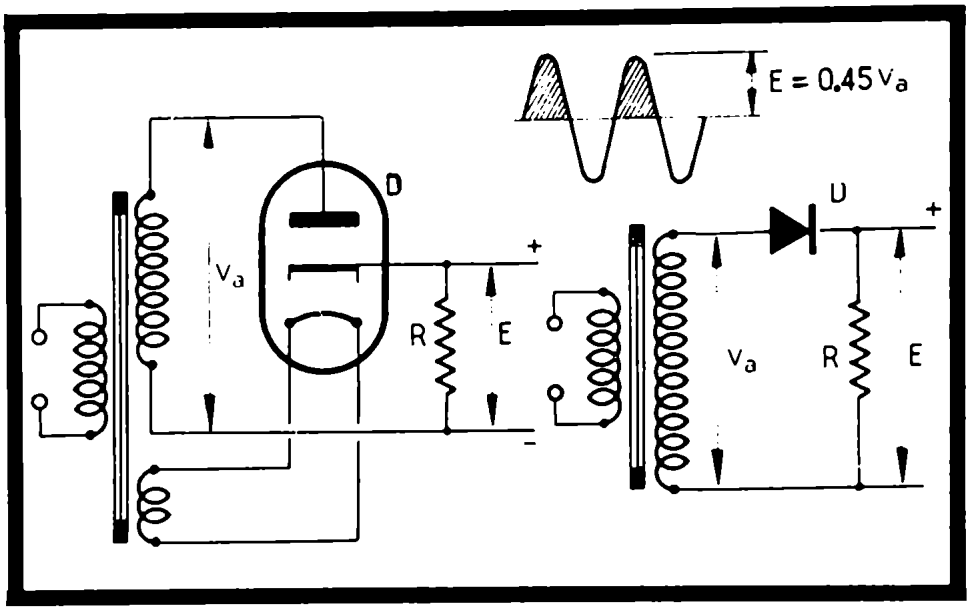


Fig. 8-1. Circuits using tube and semiconductor half-wave rectifiers.

operational life is practically unlimited (against some 2,000 to 20,000 hours for most tubes). In return, a semiconductor device generally operates nearer to its limits than does an electron tube; however, it is easily destroyed by a voltage or current transient lasting but a fraction of a second that would not harm a tube. Therefore, adequate protection is a must if the diode is not to break down with the first transient that occurs. And then, for some applications, tubes presently cost less than semiconductor devices, but this situation may change.

One-phase half-wave rectifier circuits using a thermionic and a semiconductor diode are represented in Fig. 8-1. Both diode types pass current only during the half cycles the anode is positive with respect to the cathode. The rectified voltage, E, across load resistor R, therefore, is a sequence of positive (hatched) half-cycles separated by spaces of the same duration; this is a pulsed unipolar rather than a true DC voltage. Calling Va the RMS voltage applied to the diode, the rectified voltage across a resistive or inductive load is E equals 0.45 Va.

While all types of rectifiers are damaged by an excessive current, which leads to an exaggerated heat buildup (and semiconductor devices will stand heat less than tubes), there is break-down danger during the nonconducting half-cycle if the peak inverse voltage (PIV) exceeds the rated value. Again, the safety margin of the semiconductor device, as to reverse breakdown, is narrower than that of a tube com-

parably rated. For a resistive or inductive load circuit with no input capacitor, the effective PIV equals 1.4 Va.

If a high-capacitance, C, is connected across the rectifier output (Fig. 8-2), it charges during the half cycle the rectifier conducts and restores the stored energy to load R during the half cycle the rectifier is inoperative (Fig. 8-3). The input capacitor, thus, behaves like a flywheel, and the voltage across R is no longer pulsed, but rather a mains-frequency undulating DC voltage. The "hum" component can practically be eliminated by a smoothing filter composed of choke L and a second capacitor, C1. With a filter-input capacitor, the rectified voltage is E equals 0.9 Va, or twice as much as with no input capacitor. The PIV is 2.8 Va again, twice what it was before.

Just as a heavier flywheel, due to its higher inertia, provides better speed regulation, a high-value input capacitor reduces the DC output voltage ripple. But since such a capacitor discharges less, the fraction of the cycle it actually charges is shorter than it would be with a lower-value capacitor, and for the same power delivered to the load, the charging current obviously is higher. Manufacturers specify the highest input capacitance for conservative rectifier operation. The second capacitor, of course, is not affected by this limitation.

Gas and mercury-vapor filled tubes (phanotrons) feature a low current-independent voltage drop of 10 to 15v, depending on the gas used, resulting in a higher efficiency than that realized with vacuum tubes, especially in high-current medium-voltage applications. But because of their low resistance, capacitor-input filters cannot be used with phanotron rectifiers, for a discharged high-value capacitor would draw an excessive charging current which would im-

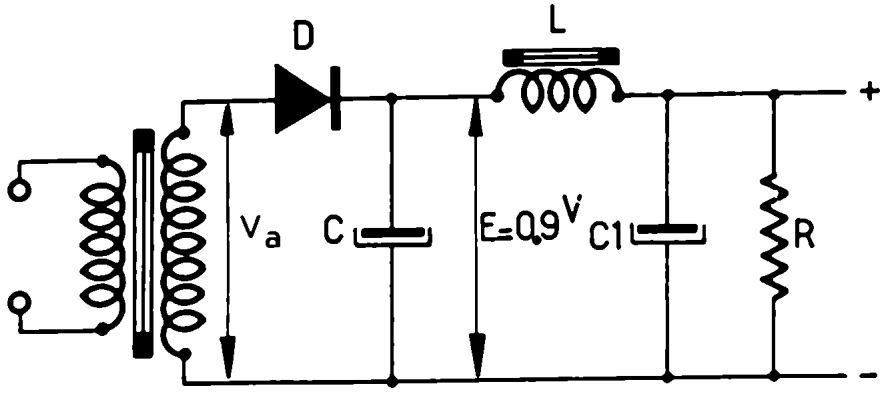


Fig. 8-2. Rectifier circuits with a capacitor input filter.

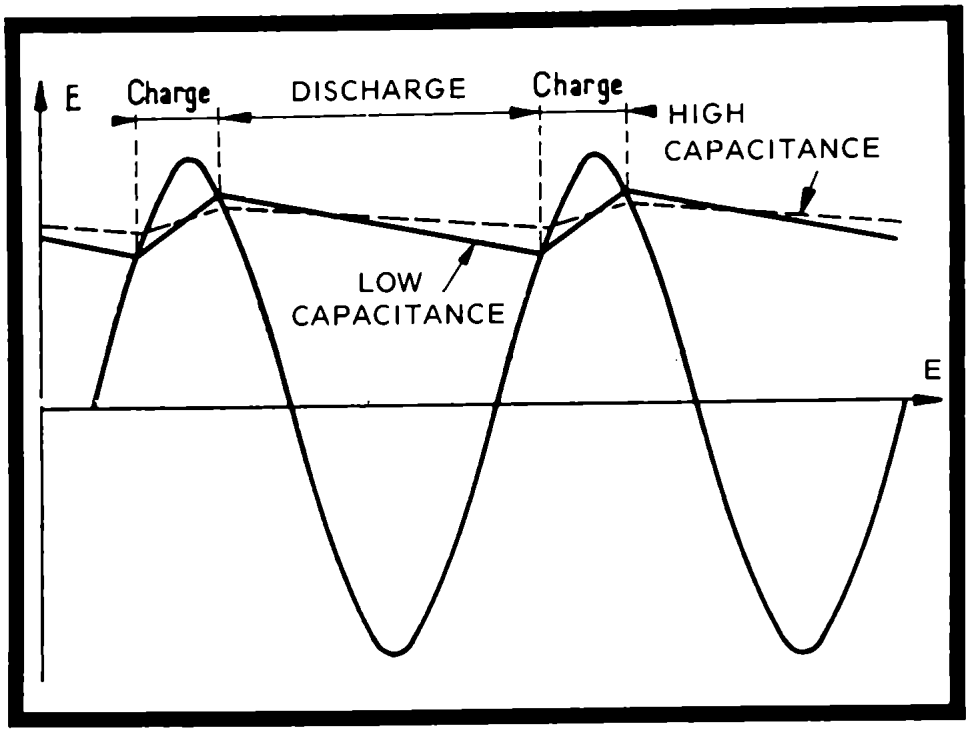


Fig. 8-3. A triangular voltage wave appears across the input filter capacitor.

mediately destroy the tube. Gas tubes, therefore, are always used with choke-input filters (Fig. 8-4). Furthermore, the AC anode voltage can be applied only after a warm-up period of several minutes, allowing for the cathode to become hot. A simple bimetallic timer will take care of this condition by delaying voltage application.

For **full-wave rectification**, a center-tapped transformer secondary, providing Va RMS volts across each half winding, feeds two diodes; there are full-wave rectifier tubes containing two anodes facing a common cathode (Fig. 8-5). Each diode passes current on alternating half-cycles, and the rectified output voltage appears to be obtained by folding up the negative half cycles of the supply voltage, for the rectified voltage now is an uninterrupted succession of positive half-cycles as shown by the hatchings. There is much less ripple than with half-waves rectification, and the ripple frequency is 120 instead of 60 Hz and is twice as easy to filter out. Since the two diodes conduct alternately, each supplies half of the load current. The rectified voltage is E equals 0.9 Va as formerly, with a PIV of 2.8 Va per rectifier.

The **bridge rectifier** (Fig. 8-6) also is a full-wave device. It uses a simple secondary winding providing an RMS voltage Va from only half the number of turns required by the full-wave vacuum tube rectifier, but four instead of two diodes are

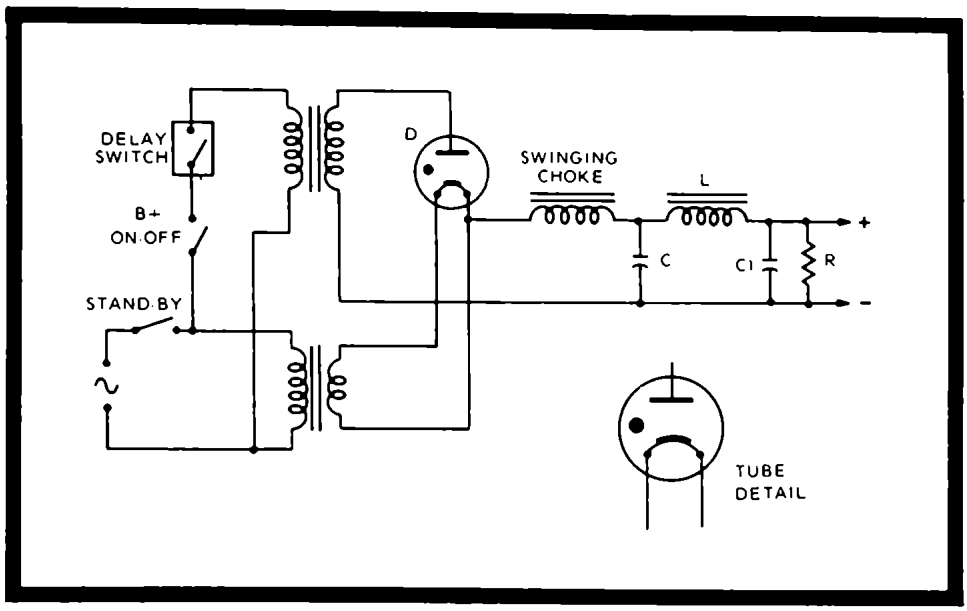


Fig. 8-4. A choke-input filter must be used with phanotron (gas-filled) rectifiers.

needed. Such rectifier bridges are marketed assembled on cooling fins and wired, ready to use. (A tube type bridge rectifier would need three separate heater windings, obviously a drawback.) Within the bridge, opposite side diode pairs conduct alternately: D1 and D2, and then D3 and D4. Since the two diodes of a pair are series-connected, the PIV per diode is only 1.4 Va, or half as much as in half-wave and full-wave conventional rectifiers. With a given diode type, the bridge rectifier will tolerate the rectification of a voltage twice as

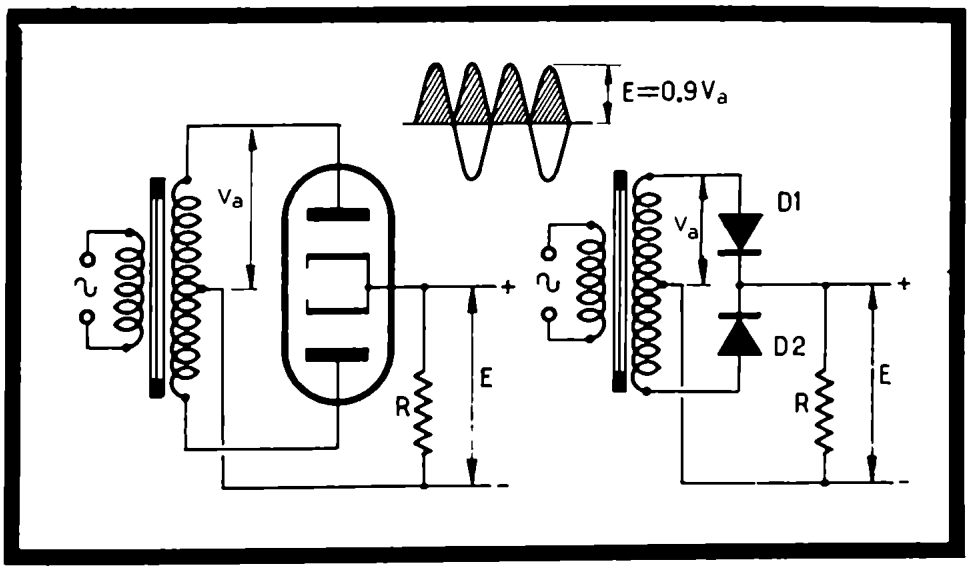


Fig. 8-5. Tube and semiconductor full-wave rectifier circuits.

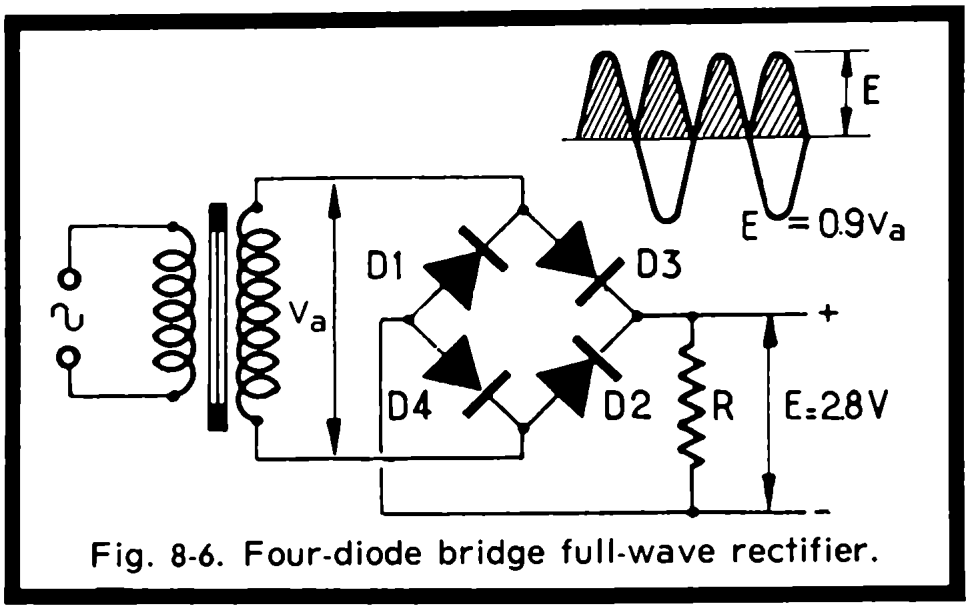


Fig. 8-6. Four-diode bridge full-wave rectifier.

high. The output, however, remains 0.9 Va and the undulating output voltage is shaped as formerly.

Voltage doublers are used when a higher DC voltage is to be obtained from an available AC supply, especially for low currents; capacitor-input filters, therefore, are mandatory. The voltage doubler in Fig. 8-7 is a full-wave device. Diode D1 charges capacitor C1 during one half cycle, and D2 charges C2 curing the next. Because C1 and C2 are series connected, the stored voltages add. Assuming negligible load current, E equals 2.8 Va. The PIV per rectifying element also is 2.8 Va.

This circuit requires a floating secondary winding. Though half-wave, the voltage doubler in Fig. 8-8 uses an AC source with one grounded terminal; hence, it is also serviceable in transformerless power supplies where a no-load DC voltage of about 615v can be obtained from the 220v mains. During the half-cycle where B is positive in regard to A, diode D2 conducts and charges C1 to 1.4 Va with the indicated polarity. During the next half-wave, D1 conducts and "sees" the input capacitor voltage of 1.4 Va in series with the AC voltage that peaks at 1.4 Va, too. Hence, the total voltage across C2 is almost 2.8 Va, again assuming negligible load current. The PIV still is 2.8 Va per diode.

For large current requirements, 3-phase rectifier systems are preferred because they don't unbalance the power network, they provide more power with components of a given rating and reduce the ripple voltage. A typical 3-phase half-wave rectifier circuit is presented in Fig. 8-9. The three rectifying elements (phanotrons, diodes or thyristors) are supplied by a delta-wye coupled transformer and the load is connected between the wye neutral and the cathodes tied

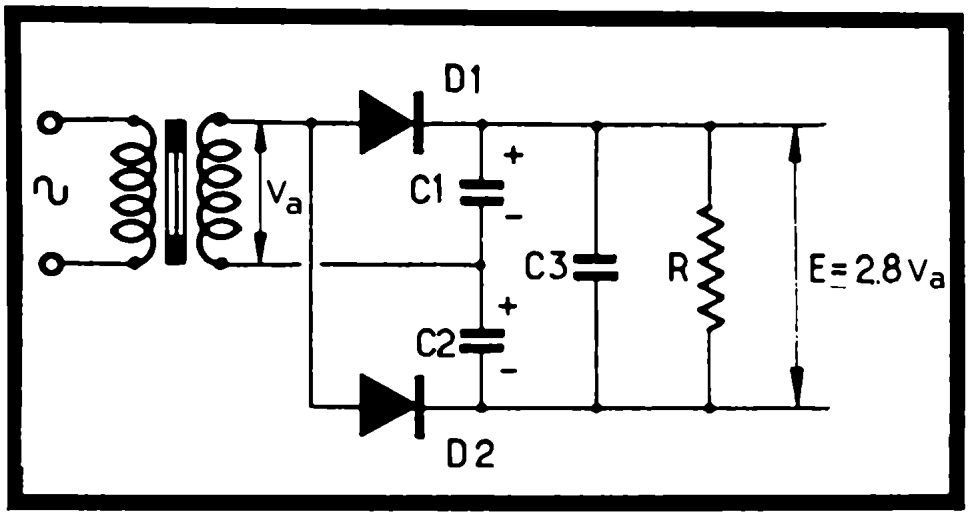


Fig. 8-7. Voltage-doubler circuit with a floating AC supply.

together. Each rectifier's conduction is offset by one third of a cycle regarding the next or the preceeding ones, so that one rectifier starts conducting before the preceding one has ceased to conduct; the resulting ripple is lower than in single-phase operation as shown by the diagram in Fig. 8-10. The comparatively high frequency (180 Hz) of this ripple minimizes the smoothing (if any). And then, each device has to carry one third of the load current. Here, E equals 1.17 Va, and the PIV is 2.45 Va per device.

Still less ripple is obtained with the **3-phase full-wave rectifier bridge** using six phanotrons, rectifiers or thyristors

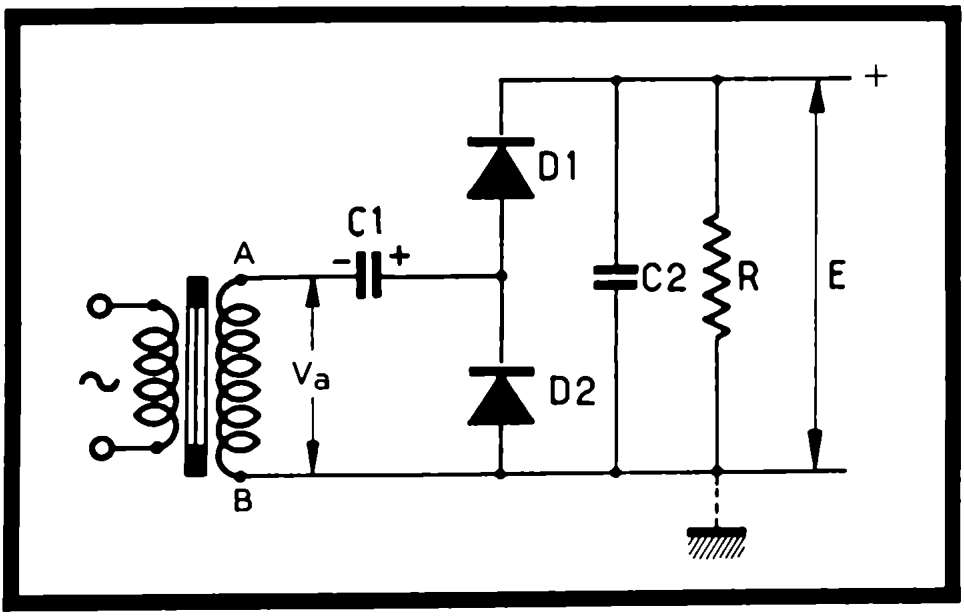


Fig. 8-8. Voltage-doubler circuit with a grounded AC supply.

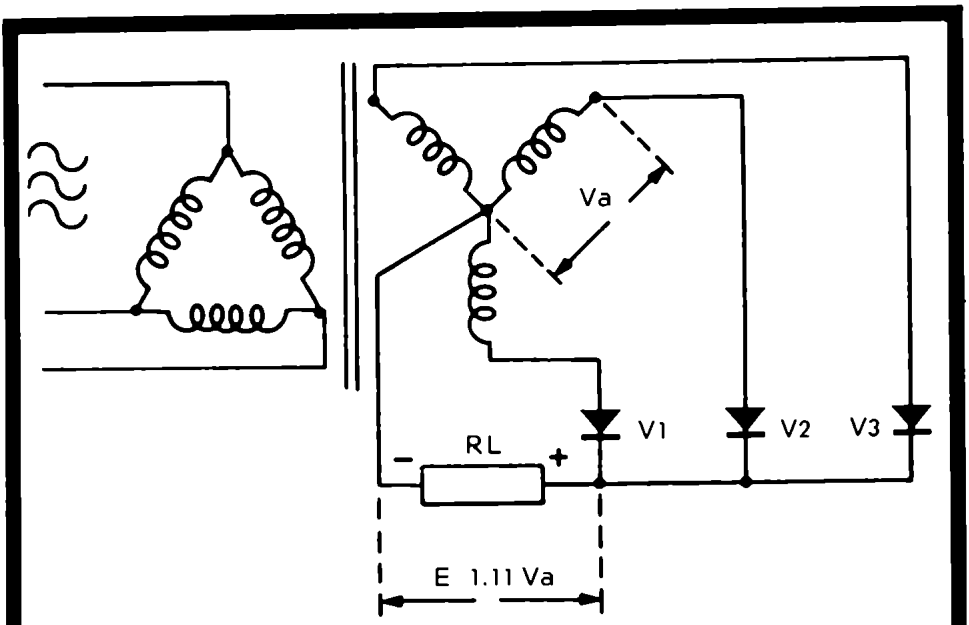


Fig. 8-9. Three-phase half-wave rectifier circuit.

(Fig. 8-11). Such semiconductor rectifiers assembled on cooling fins and wired are readily available. Two devices conduct at the same time, one of the upper row (V1, for instance) and, successively, the two lower row devices not directly beneath (V5 and V6). Hence, there are six operating sequences during one cycle: V1-V6, V1-V5, V3-V5, V3-V4, V2-V4 and V2-V6. Each rectifier conducts during one third of a cycle, and a new component starts conducting every sixth of a cycle. This makes for very low ripple of still higher frequency (360 Hz for a 60-Hz line frequency). The rectified voltage now is E equals 2.34 Va, again with a PIV of 2.45 Va per rectifying element.

CONTROLLED RECTIFIERS

The controlled rectifier allows for varying the rectified power within wide limits while draining only a corresponding amount of power (neglecting losses). At least for power applications, a controlled rectifier, therefore, is more economical to operate than an ordinary rectifier that converts power excesses into heat. Furthermore, a perfectly smooth variation can be obtained, while the rectifier supply voltage variation by a tapped transformer is discontinuous.

Controlled rectifiers are thyratrons or thyristors that can only conduct during one half-cycle, proper control electrode (grid, gate) bias provided. Once a controlled rectifier is fixed,

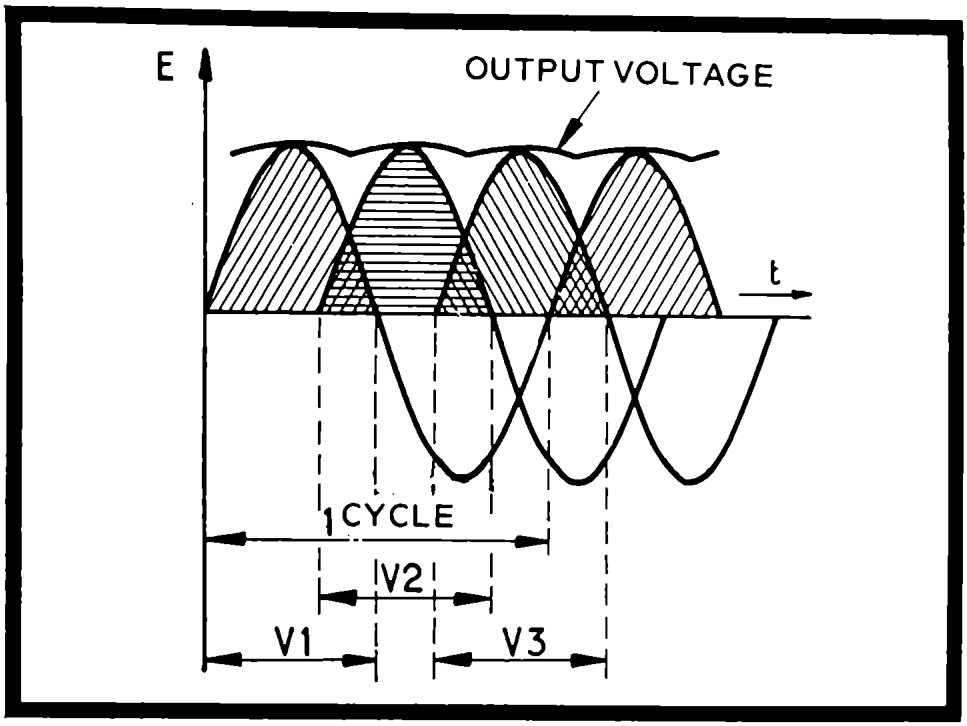


Fig. 8-10. The ripple content of a 3-phase rectifier output is low.

the grid or gate looses control for the half-cycle concerned at the end of which the controlled rectifier de-energizes anyway.

The rectified power can be varied by varying the conducting angle of the device; i.e., the fraction of the cycle the rectifier conducts. By varying the DC bias voltage, the so-called "vertical control," the operating point can be shifted only between the beginning (A) and the peak (B) of the positive half-wave (Fig. 8-12), resulting in a limited variation of the rectified power. A preferred approach, called

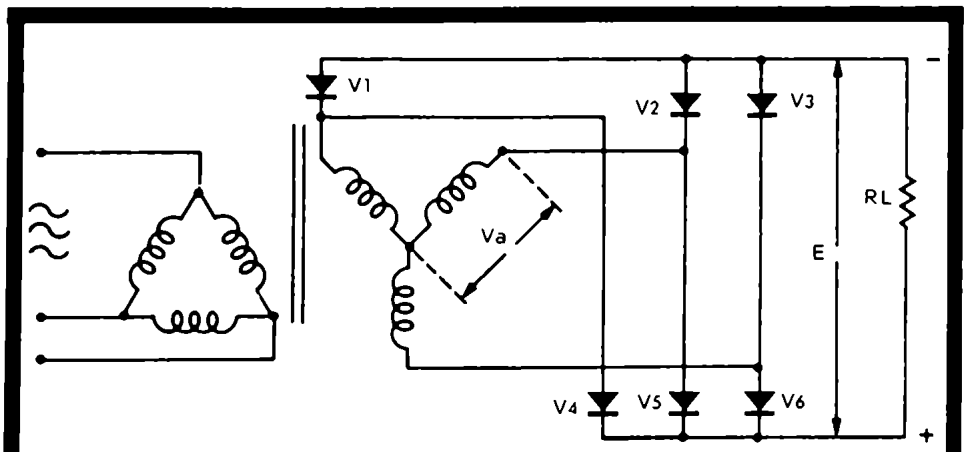


Fig. 8-11. Full-wave 3-phase bridge rectifier circuit.

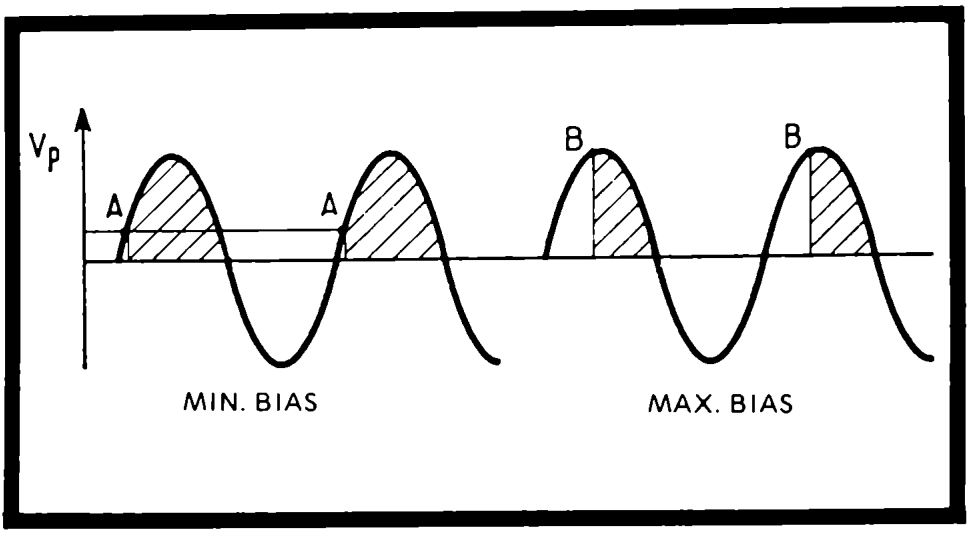


Fig. 8-12. Vertical control bias provides only a limited variation of the controlled rectifier output.

"horizontal control," shifts the striking point along the conducting half-cycle by a phase-variable AC voltage (Fig. 8-13). This allows the rectifier output to be varied from almost nothing to nearly full half-wave power. The oscillograms in Figs. 8-14 to 8-17 illustrate the gradual reduction of the controlled rectifier conduction angle from nearly all of the half-wave to a small portion of it. The amount of power actually

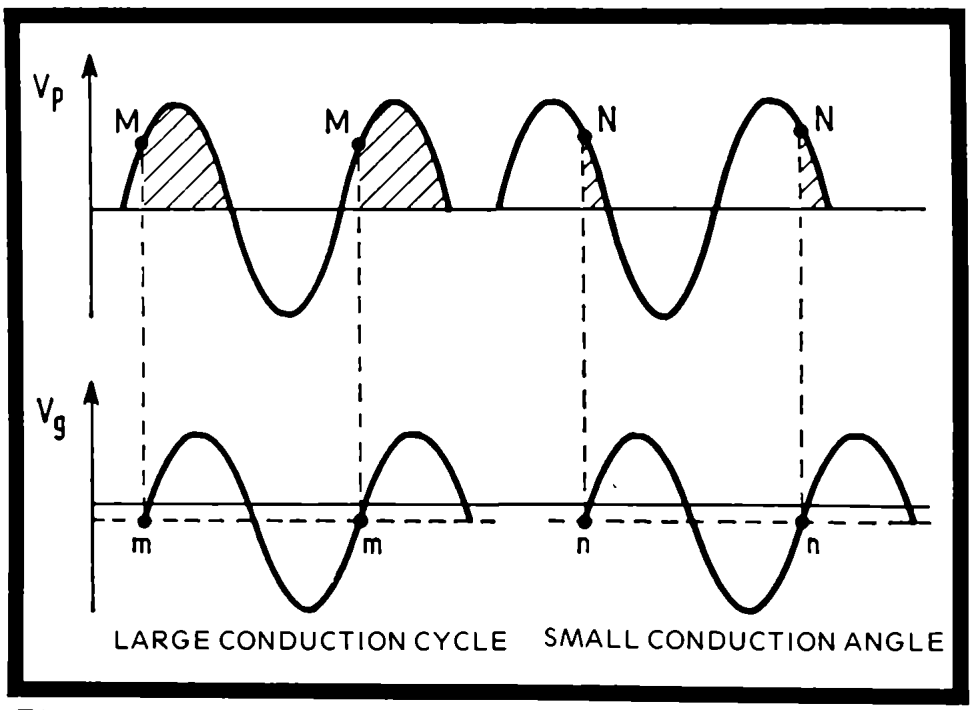


Fig. 8-13. Horizontal control allows continuous variation of the controlled rectifier output.

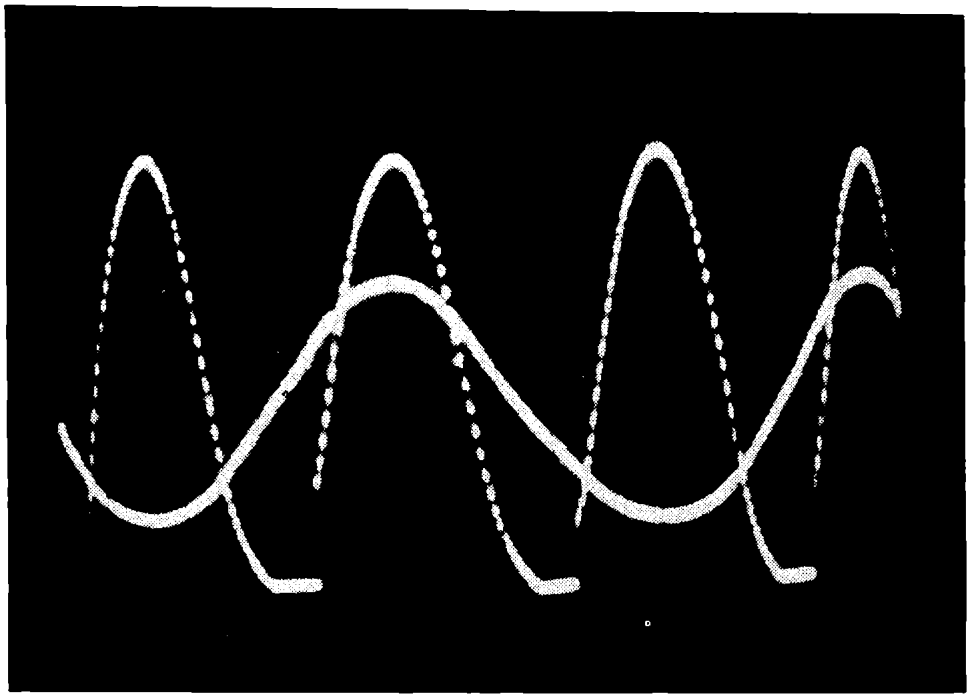


Fig. 8-14. Oscilloscopes showing almost full controlled rectifier output.

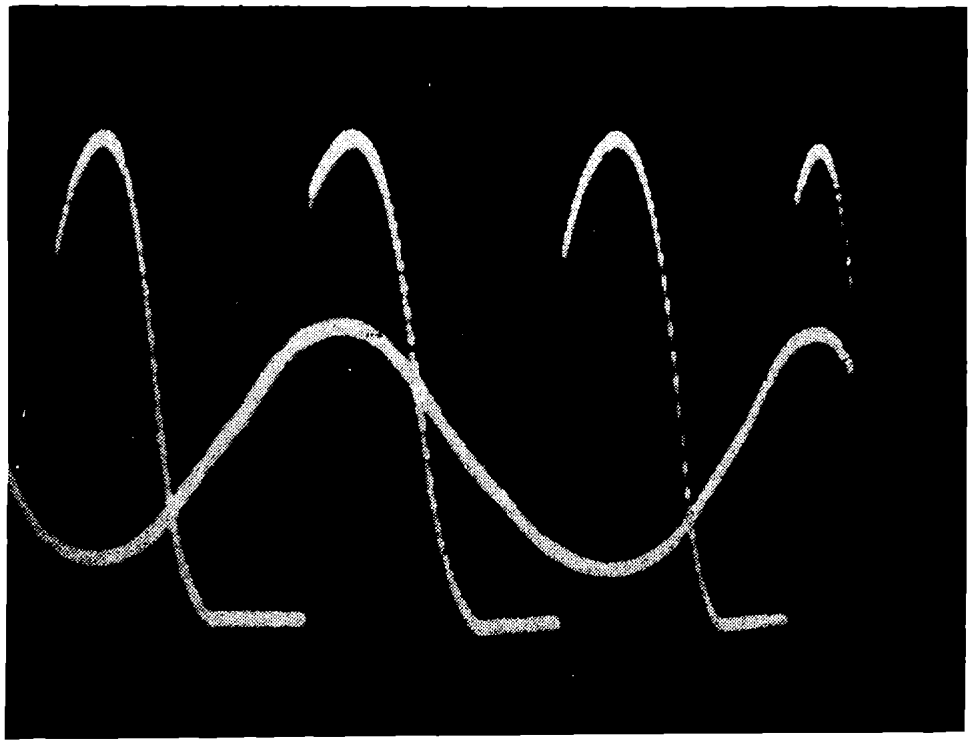


Fig. 8-15. Oscillogram showing slightly more than half-power output.

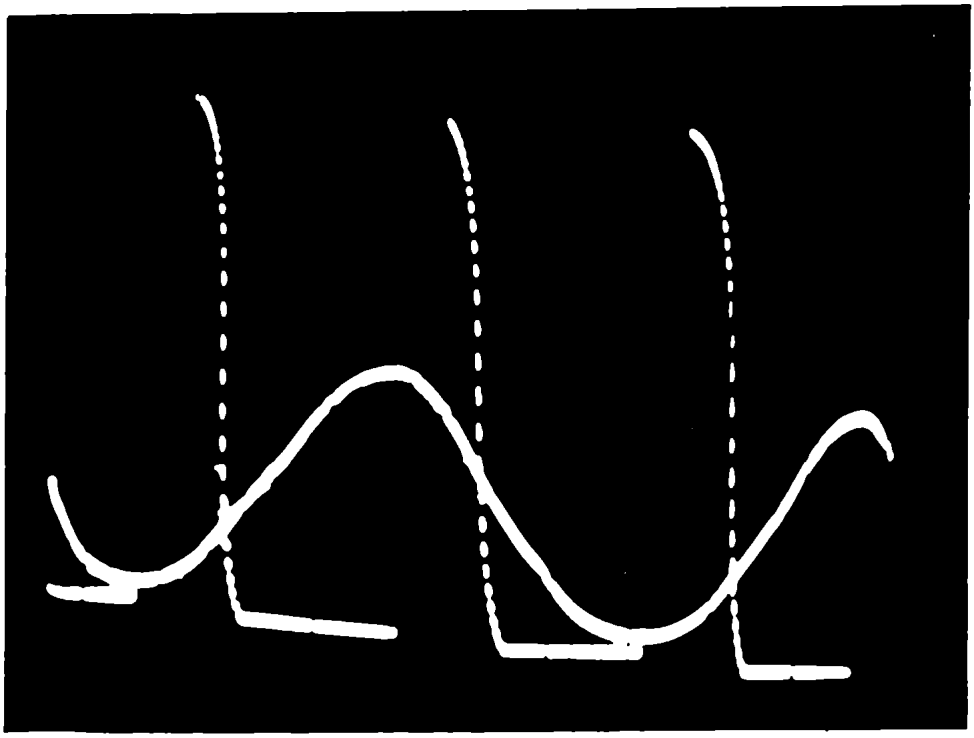


Fig. 8-16. Oscillogram showing exactly half-power output.

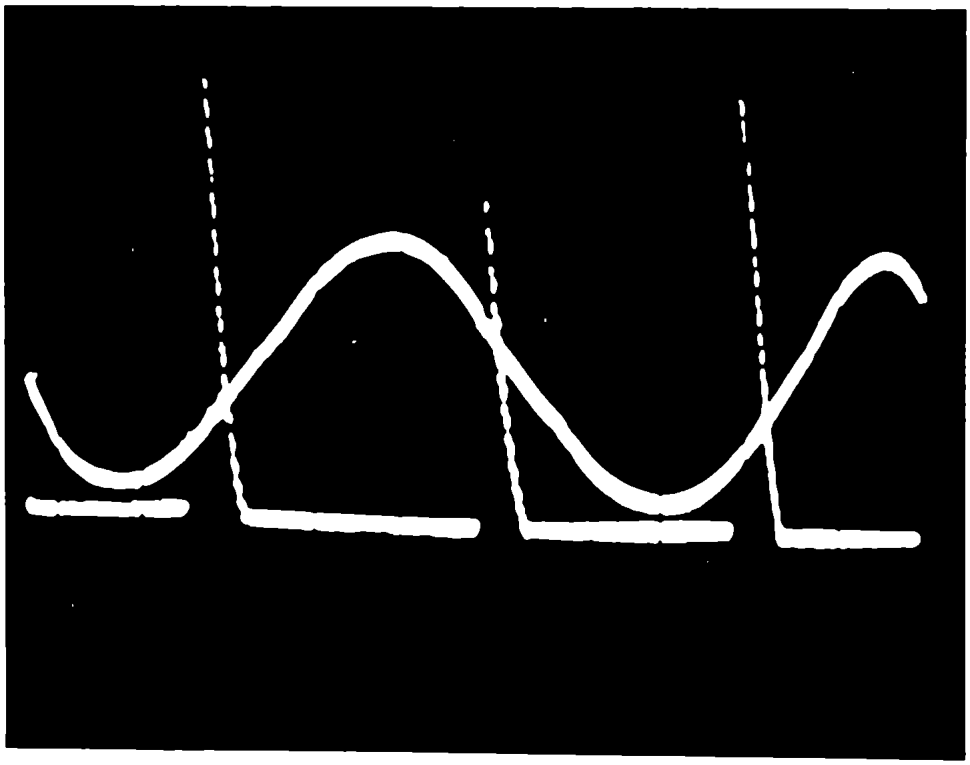


Fig. 8-17. Oscillogram showing less than half-power output.

rectified and delivered to the load is proportional to the area located between the curve and the abscissa. (The superimposed sine wave is the anode voltage of one of the two thyratrons of a full-wave controlled rectifier.)

A simple one-thyristor half-wave controlled rectifier circuit using a phase bridge for horizontal control is shown in Fig. 8-18. The conduction angle of thyristor SCR is varied by potentiometer R, forming a bridge circuit with capacitor C and the center-tapped transformer secondary. Load RL passes a pulsed current starting at a point which is adjustable along the conducting half-cycle, ending with the end of this half-cycle. Diode D short-circuits the control signal during the negative half-cycle, and zener diode, DZ, limits the amplitude of the signal during the conducting half-cycle. This circuit is serviceable as a variable-speed control for fractional horsepower mains-fed series and "universal" motors.

In the full-wave phase-controlled rectifier bridge in Fig. 8-19, thyristors SCR1 and SCR2 are associated with diodes D1 and D2 to control the current in load RL; SCR1 and D2 conduct during one half-cycle, and SCR2 and D1 during the other. The same phase bridge, RC, controls both diodes; D3 and D4 short-circuit the control signal during the nonconducting half-cycle of the corresponding SCR. For reliable operation, the peak value of control voltage Vc must exceed 25v.

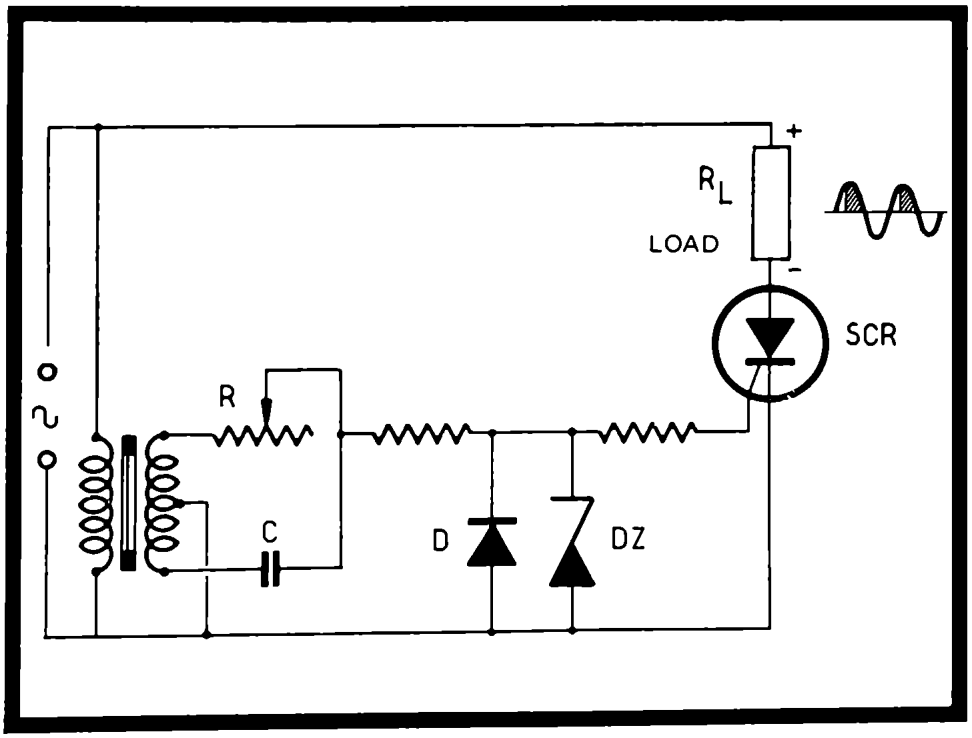


Fig. 8-18. Schematic of a phase-controlled thyristor half-wave rectifier.

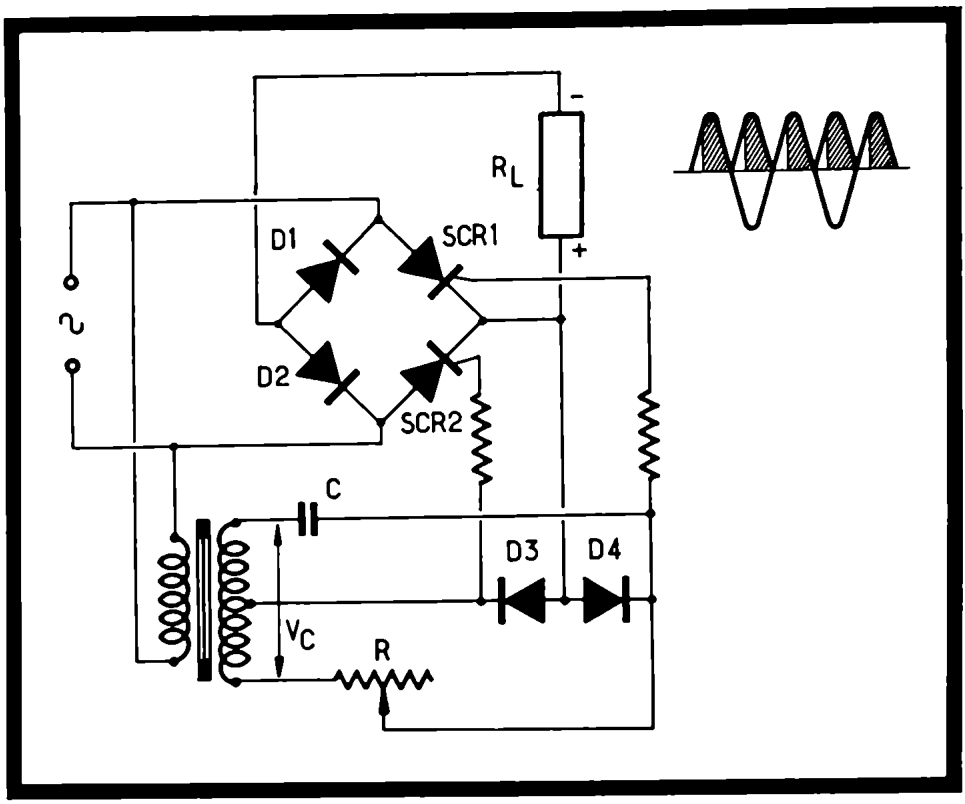


Fig. 8-19. Full-wave phase-controlled thyristor rectifier bridge circuit.

While thyristor firing by a phase-shifted sine wave is generally satisfactory, pulse control is more accurate and flexible. Unijunction transistors are a natural for SCR control, since they deliver just the control pulse type required. In the full-wave load current control in Fig. 8-20, SCR1 and SCR2 are triggered by a single unijunction transistor, UJT. Diodes D1 and D2 provide full-wave rectification of the transformer voltage. (If this AC voltage is less than about 100 RMS volts, a separate winding is to be used to provide sufficient operating voltage to the UJT.) Zener diode DZ, series-connected with resistor R1, clips the **nonsmoothed** half-cycles to shape them approximately flat-topped; this is necessary to synchronize the UJT operation with the mains frequency in order to fire the SCRs at the right moment. With no capacitor in the UJT supply circuit, the voltage falls to zero twice each cycle at the sine-wave zero-crossing points. The UJT striking delay with respect to the beginning of the half-cycle (corresponding to the sine-wave zero-crossing) is adjusted by variable resistor R and capacitor C. The UJT delivers a trigger pulse at every half-cycle. All these pulses are simultaneously injected into both SCR gates, but only the SCR fed with the conduction half-cycle is responsive.

A full-wave 3-phase controlled rectifier diagram is shown in Fig. 8-21. The rectifier bridge contains three thyristors SCR1, SCR2 and SCR3, associated with three diodes D1, D2 and D3 of similar ratings, and a single UJT for triggering the SCRs. This comparatively simple circuit allows for output control between 25 and 100 percent of full rectified voltage; there are more sophisticated circuits for control between 0 and 100 percent, but the added complication may not be worthwhile. Diodes D4, D5 and D6 rectify the trigger generator supply voltage clipped by zener diode DZ, and the SCR triggering pulse is produced by UJT1 as explained above.

In three-phase supplies, the SCR firing angle is to be variable over 120 degrees or one third of a cycle. If 120 degrees is exceeded, the SCR of the next phase is fired at the beginning of its working half cycle, and the rectified voltage, up to then rather low, bluntly jumps to its top value. This drawback is avoided by UJT2 and T. UJT2 is a relaxation oscillator independent of UJT1, but since it is fed the same clipped voltage, timing starts at the same instant for both devices. Resistor R2

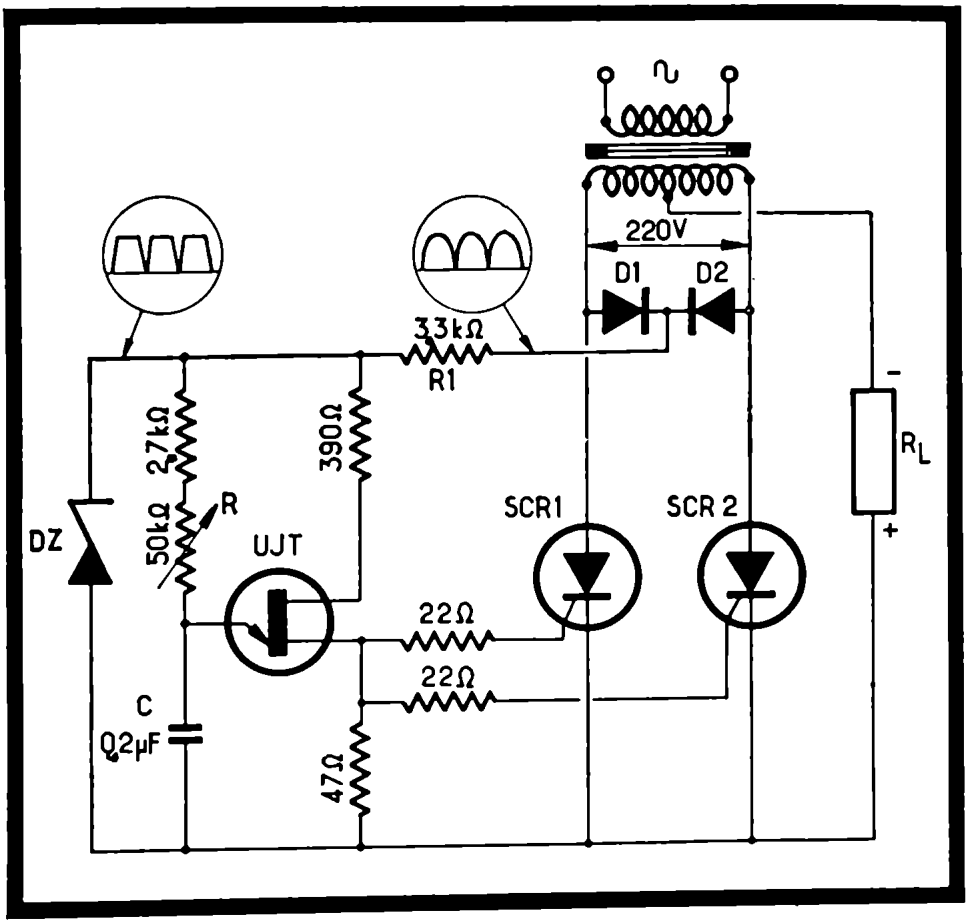


Fig. 8-20. Unijunction-transistor triggered full-wave thyristor rectifier circuit.

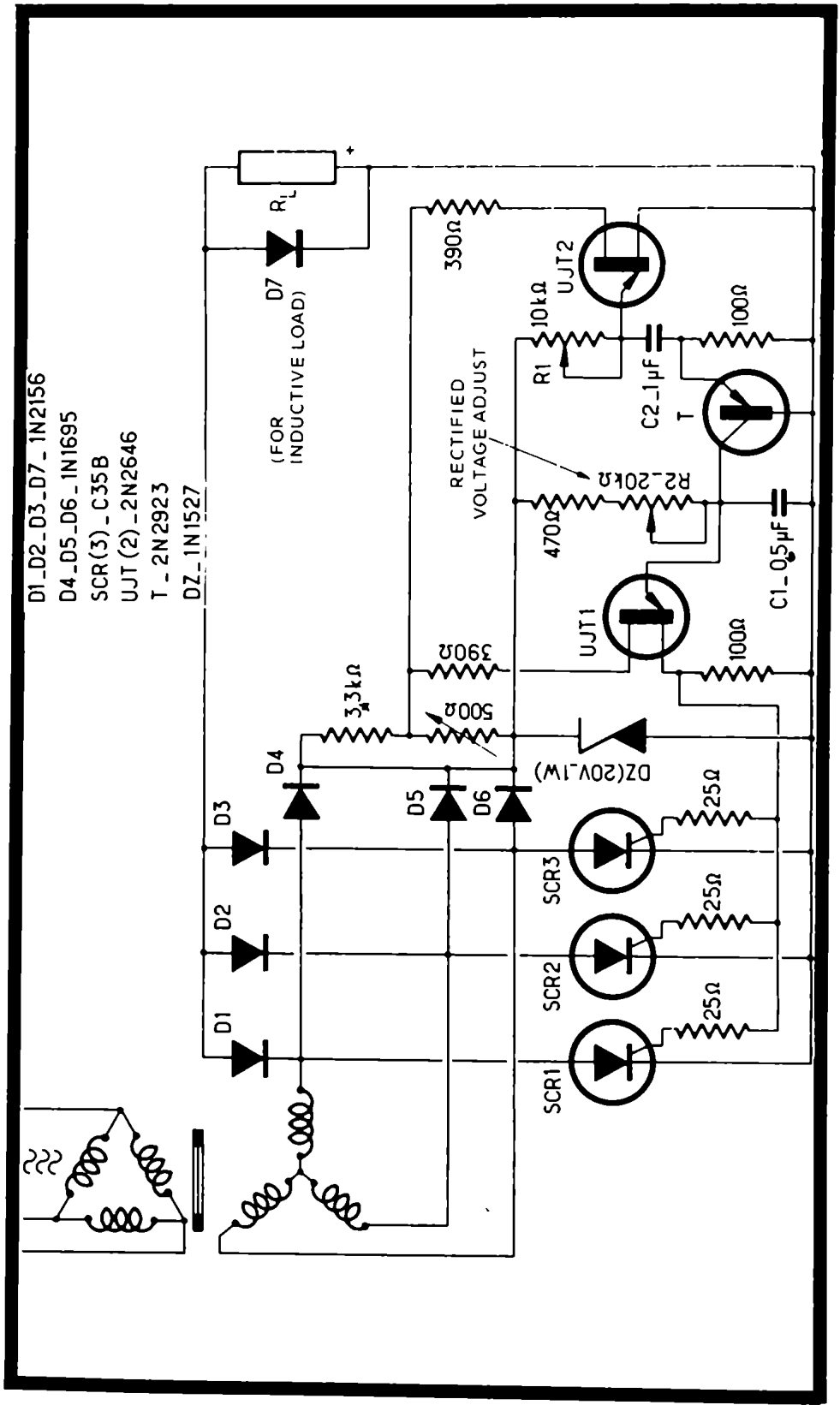


Fig. 8-21. In this circuit, one UJT triggers a 3-phase thyristor rectifier bridge.

is adjusted to limit the firing angle of UJT2 to slightly less than 120 degrees. If the firing angle of UJT1 exceeds 120 degrees at any instant, UJT2 strikes and discharges C2 through the emitter-base junction of T. This saturates the transistor and that discharges capacitor C1 in the UJT1 gate circuit. Thus, there is no SCR triggering pulse, hence no rectification.

In all these circuits, the variable resistor, fixing the striking angle of an SCR or UJT, can be replaced with a signal-controlled transistor, or even a suitable resistive transducer.

REGULATED POWER SUPPLIES

For consistent and reliable operation of any equipment, its supply voltage must remain constant. A regulated power supply provides such a constant voltage, independent of fluctuations occuring **downstream** (load variations) as well as **upstream** (line voltage instability). While the latter evidently needs regulation, the influence of load variations must be considered also.

A voltage source can be represented by a generator of negligible internal resistance, series-connected with a resistor Ri, the (internal) generator resistance (Fig. 8-22). Calling E the emf of this generator, output voltage Vo across load resistor RL is Vo equals E - RiI, where I is the load current. While Vo equals E in the no-load condition (I equals 0), the voltage drop RiI increases with I, so that Vo falls when RL draws a higher current. To make Vo independent of load current variations, the generator voltage drop v equals RiI must be made negligible; hence, Ri must be much smaller than RL. Regulated power supplies, indeed, feature an Ri that is a fraction of an ohm in low-voltage supplies and several ohms in B+ supplies. Such a low resistance appearing across the power supply terminals effectively reduces unwanted amplifier interstage coupling through the common source impedance, thus avoiding uncontrolled feedback.

If a supply of several hundred volts is to be regulated, the long-known tube-type circuits may still have an edge. The

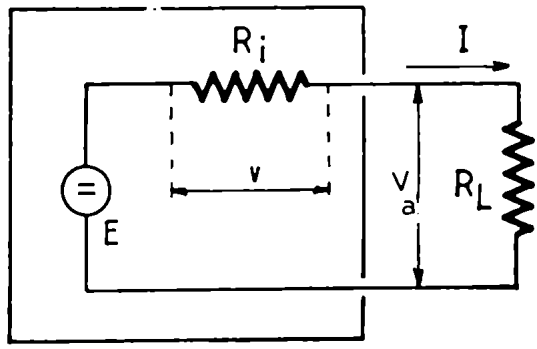


Fig. 8-22. In a power supply circuit, source resistance Ri is in series with the output.

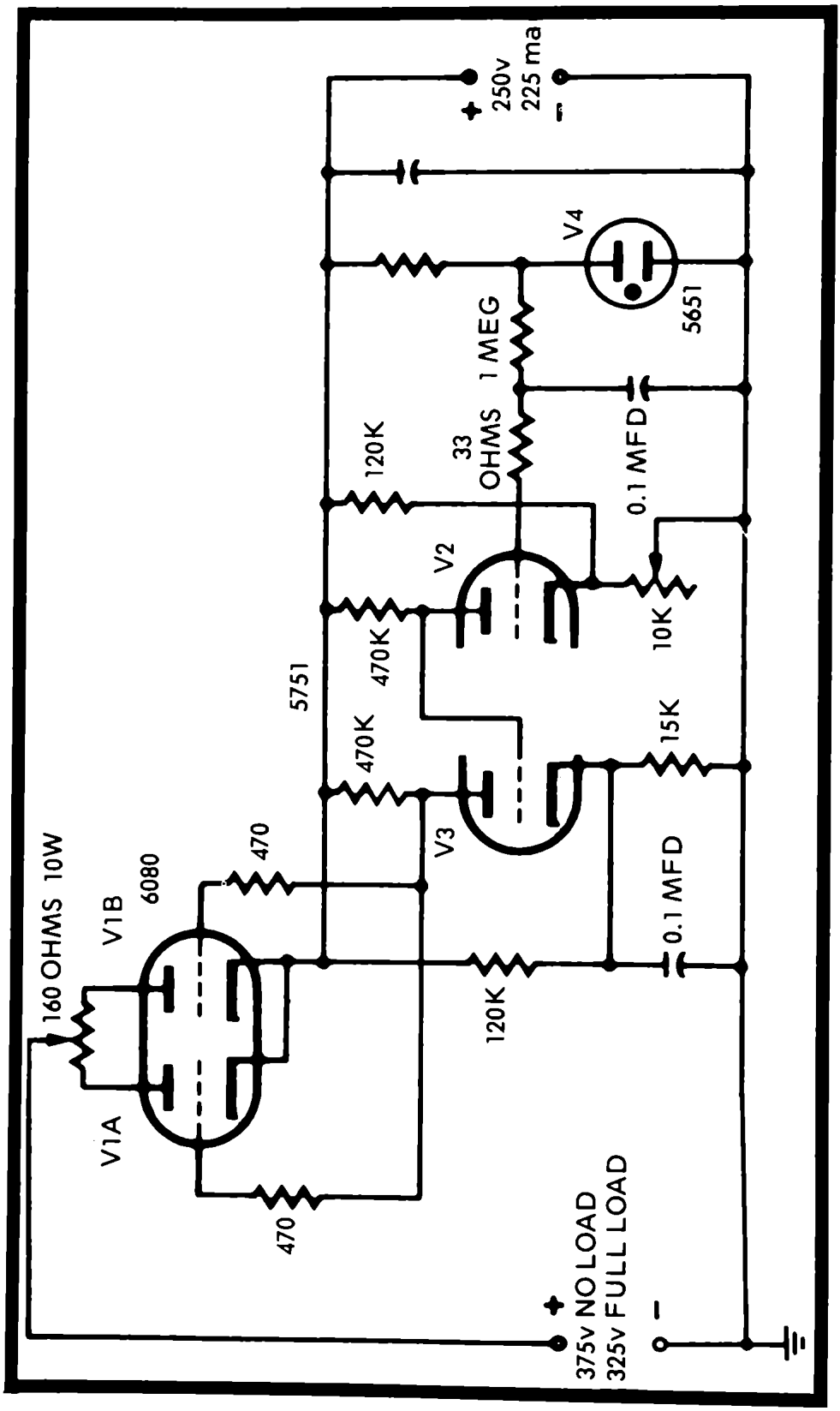


Fig. 8-23. This tube-type regulated power supply circuit delivers 250v at 225 ma.

regulated voltage supply in Fig. 8-23 delivers 250v with a full-load current of 225 ma. Since its source resistance is only about one ohm, output voltage Vo varies by less than 0.2v for a load current varying from 0 to 225 ma. This is obtained by an error signal amplifier composed of cascaded triodes V2 and V3, featuring an overall gain exceeding 1,000. The series regulator tube (6080) is a special twin triode featuring an internal resistance of only 250 ohms per triode. For a line voltage fluctuation of plus or minus 10 percent, Vo varies by less than 0.1v.

The typical transistor-regulated low-power voltage supply presented in Chapter 1 (Fig. 1-8) illustrates the closed-loop regulation principle. Transistor T1 is used as a controlled series resistance, with a voltage drop (dv) made variable to obtain constant Vo. T2 compares output-proportional voltage V1 with zener-diode DZ stabilized reference voltage Vr and injects into the T1 base a current adequate to maintain Vo at the specified (preset) value.

In another low-voltage power supply (Fig. 8-24), emitter-coupled transistors T2 and T3 form a differential amplifier, with the bases tied to zener-stabilized reference voltage Vr and output-proportional divider voltage V1, respectively. The voltage comparator thus obtained applies a current to the base of series transistor T1 so that Vo remains constant. The low-power supply (60 ma) features a source resistance of 0.2 ohms.

If such a variable power supply is adjusted to deliver a rather low voltage with a high current, the series transistor must dissipate a considerable amount of power, requiring a higher-rated power transistor assembled on a large-size cooling fin. However, transistors of lower rated power can be

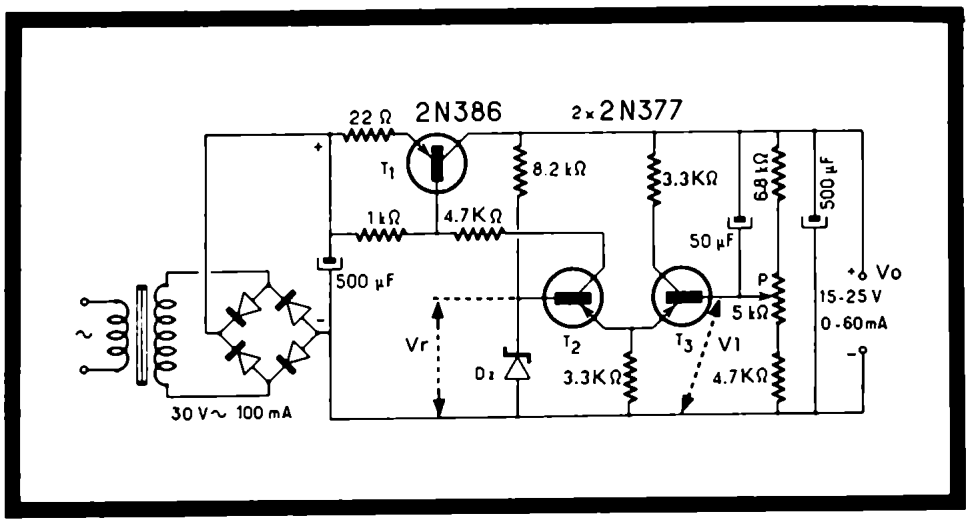


Fig. 8-24. Transistorized low-voltage regulated power supply circuit.

used if the transistor is pulse-controlled by an oscillator in a way to vary its mark-space ratio according to the error signal. Thus, the transistor delivers only the current necessary to maintain the required output level. Such a pulse-operated power supply generates strong interference, and the necessary shielding may prove costlier than the straightforward solution using an amply-rated transistor.

CONSTANT-VOLTAGE AND CONSTANT-CURRENT SUPPLIES

Besides constant-voltage supplies, there are constant-current sources that allow for linear capacitor charging and for operating current-controlled devices such as transistors. Now it's the current and not the voltage that must be load-independent. This can be obtained by making Ri large regarding RL by means of a high-value series resistor. While this approach is practical only if small currents such as a transistor base current are to be regulated, any current can be conveniently regulated by electronic means.

Constant-voltage and constant-current supplies are not very different in design. In a constant-voltage supply (Fig. 8-25A), the error signal which controls series transistor T1 is sampled across the overall output or part of it; hence, it is proportional to Vo. This signal will control T1 to maintain a constant voltage across load resistor RL. In a constant-current supply Fig. 8-25B, the error signal is sampled across a small-value resistor (r) traversed by the load current, and voltage drop rIo is amplified and applied to T1 to control its resistance so that load current Io remains constant.

Transistorized regulated power supplies have matured to a degree of sophistication that tube types never attained. Besides excellent regulation characteristics, switching from the constant-voltage to the constant-current mode is often provided, as well as current limiting according to a program and quick-response overload protection (essential to protect regulator transistors against breakdown).

LINE-VOLTAGE REGULATION

Widely fluctuating line power calls for line voltage regulation; this rather than output DC voltage regulation is the most economical approach if a number of electronic systems are to be operated. Such an AC regulator obviously reduces only fluctuations upstream, and a complementary DC (downstream) regulator may be required in addition.

To regulate powers of at least several kilowatts, variable autotransformers are sometimes used, with voltage

regulation achieved by a small slider-driving motor (Fig. 8-26). The rectified output voltage is compared with a reference voltage. The resulting error signal is amplified and energizes the motor to rotate in one direction or the other in order to maintain a constant output. Such a regulator introduces no distortion of its own, it stands reactive loads and features an efficiency of some 97 percent. Its comparatively slow response

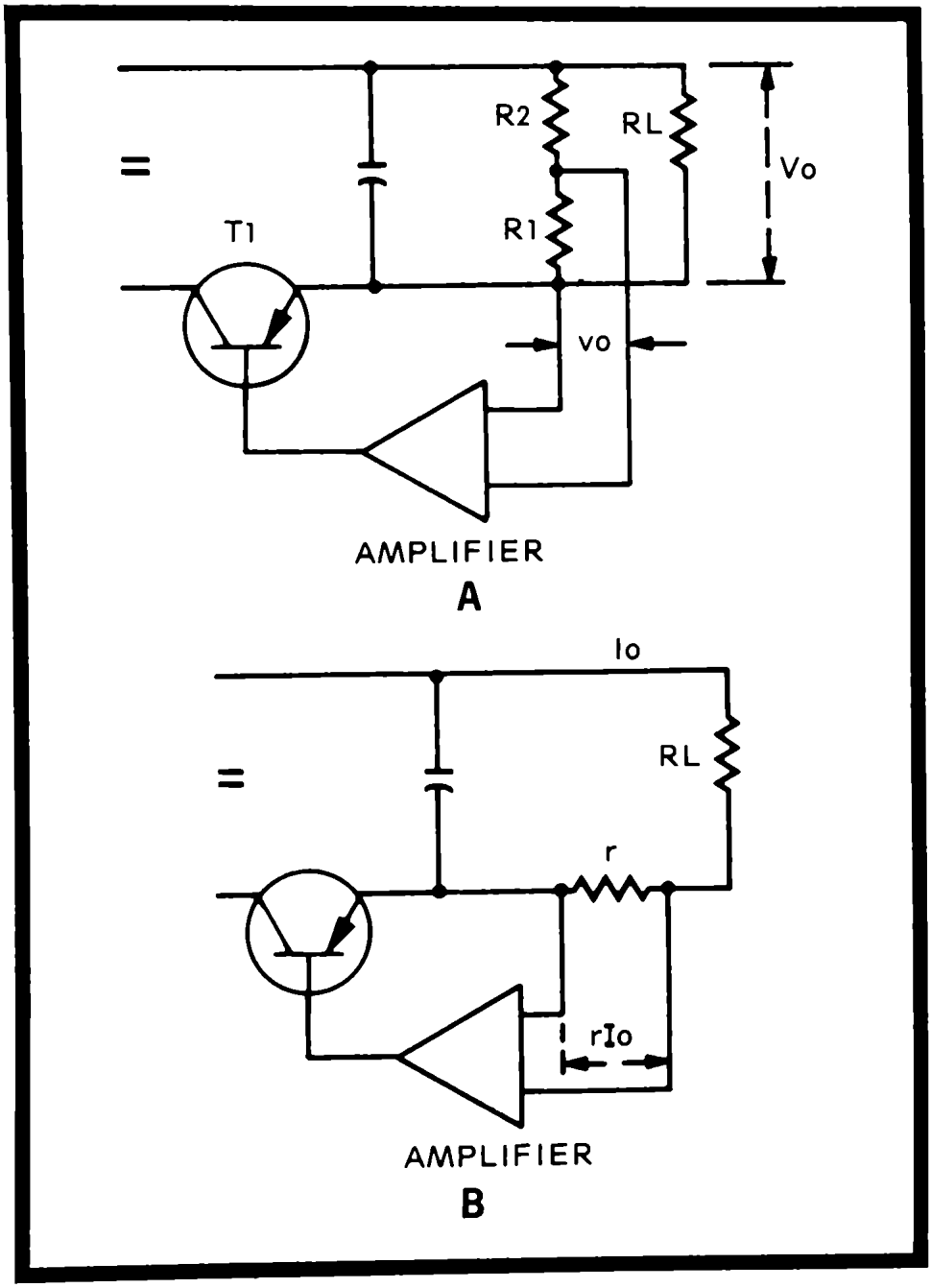


Fig. 8-25. These circuits represent typical constant-voltage (A) and constant-current (B) power supplies.

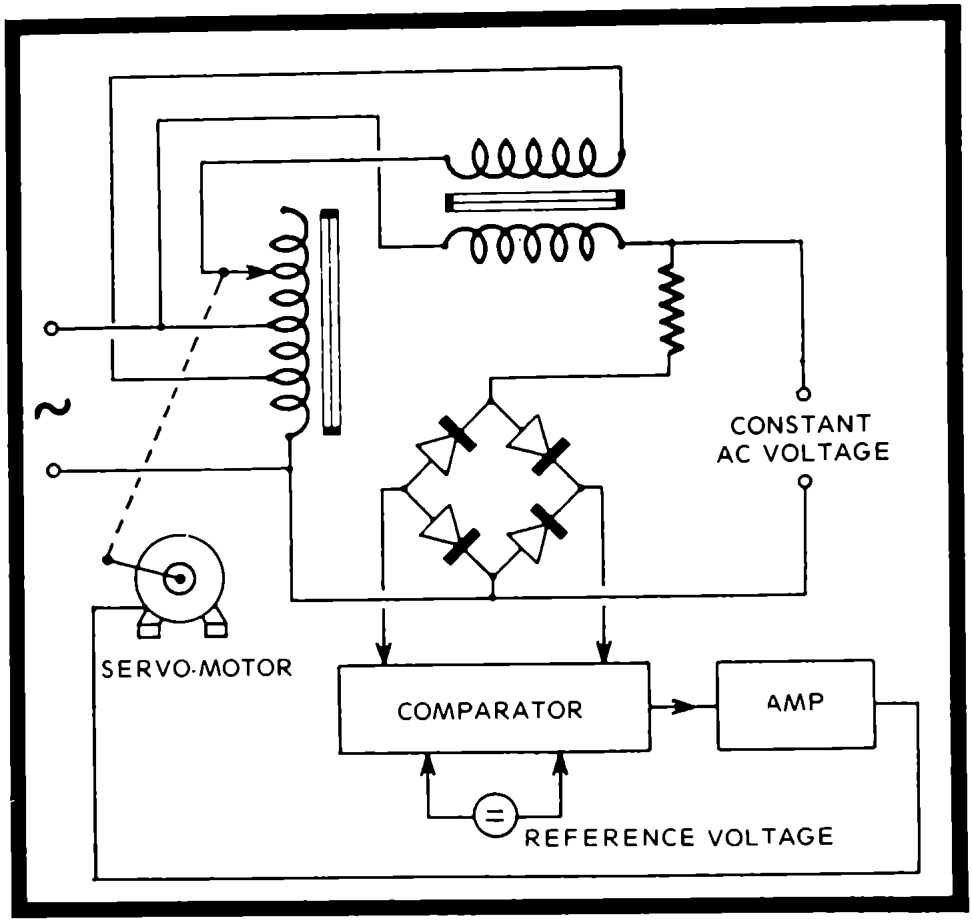


Fig. 8-26. Schematic of a motor-driven AC voltage regulator.

(0.2 second per volt), however, does not compensate quick fluctuations.

Excellent regulation characteristics are obtained with electronically-controlled saturable-reactor regulators such as the Sorensen constant AC-voltage supply whose (simplified) diagram is shown in Fig. 8-27. Line-voltage Vi is applied to part of autotransformer winding T1, which is series connected with saturable reactor load winding LL. A DC current (Ip) traversing control winding Lc more or less saturates the magnetic core, thereby reducing reactance LL and increasing the primary current of T1, hence output voltage Vo.

The voltage-sensitive transducer used here is a heater diode (V1) specially manufactured for this instrument and operated in the saturated condition. Its internal resistance is directly related to heater voltage Vh and practically independent of its anode voltage. Secondary L1 of transformer T2 (with the primary is connected across the output terminals) supplies a reference voltage to the diode heater, while L2 provides the anode voltage. With resistors R1, R2 and R3

the diode forms a Wheatstone bridge. An increase of the regulated output voltage Vo—and of Vh by the same token—reduces diode resistance. This drives the grid of power tetrode V2 more negative, lowering plate current Ip. The reactor becomes less saturated, decreasing Vo.

To compensate for output fluctuations brought about by load current variations, transformer T3 samples the load-current proportional voltage drop across series resistor r which is also applied to the diode heater. This regulator is said to feature an output voltage variation of only plus or minus 0.1 percent for a line voltage fluctuation of 15 percent, and a similar stability regarding load current variations. In recent models, a transistor circuit probably supersedes tetrode V2, and a zener diode or other semiconductor device possibly replaces heater diode V1.

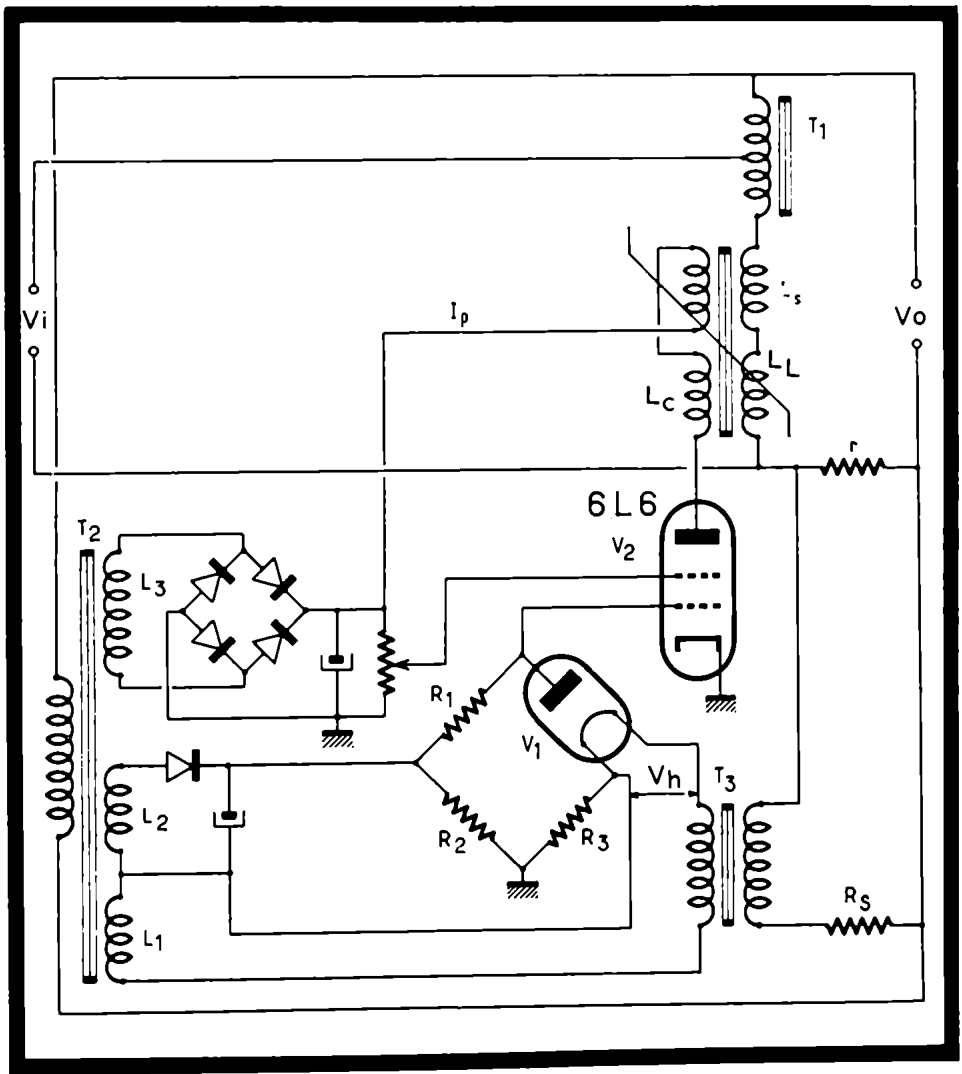


Fig. 8-27. Electronically controlled saturable reactor AC voltage regulator.

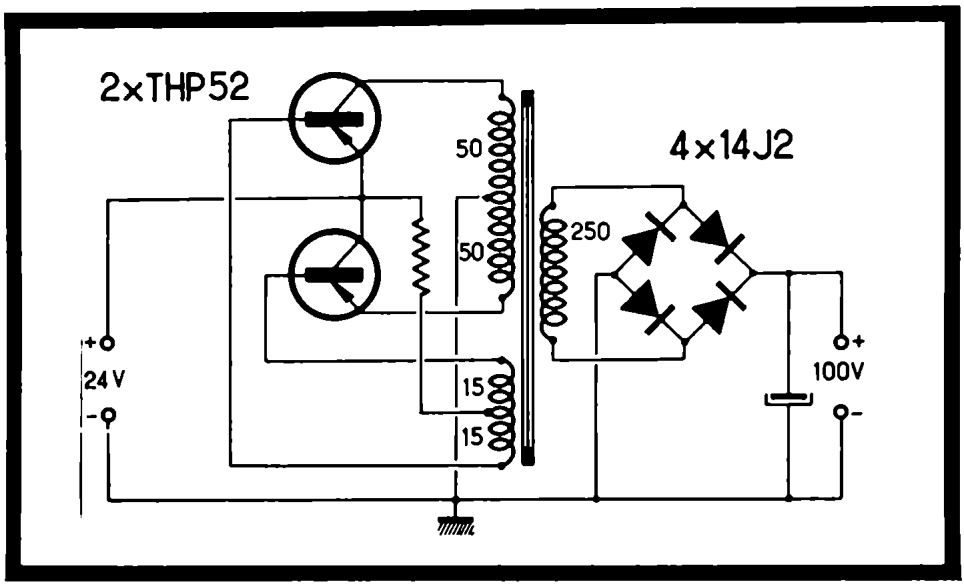


Fig. 8-28. Two-transistor self-excited DC converter using the push-pull configuration.

INVERTERS AND CONVERTERS

Before we begin this discussion, we must clarify the terminology used for converting electric energy of one form into another to avoid confusion:

Transformer	transforms AC into AC
Rectifier	transforms AC into DC
Inverter	transforms DC into AC
Converter	transforms DC into DC

The simplest energy converter type is the transformer; unfortunately, its use is limited to AC-to-AC conversion. (Rectifiers were discussed at the beginning of this chapter.) If the source available is a DC type (a battery, for instance), it can be turned to AC by means of an inverter. Additional equipment may be required if a close-tolerance frequency stability is specified. A (DC) converter is something like a DC transformer which changes one DC voltage into another; this is done by rectifying the output of an inverter. Since no AC appears in the output, no frequency regulation is to be considered and the operating frequency is chosen for the best conversion efficiency.

Any LC oscillator can be considered an inverter, as well as a converter if its AC output is rectified. Using sine waves, the conversion efficiency, however, will be poor and the transistor will generate only a small amount of power. On the other hand, if the transistor is operated as a switching device (Chapter 3),

efficiencies exceeding 70 percent are obtained. Indeed, a nonconducting transistor features a rather low residual current, and in the saturated state, its resistance is quite low, 1 ohm or so.

For the transistor to tip over as quickly as possible from the blocked state to saturation and vice versa, the core material of the inverter or converter transformer features a square hysteresis loop (such as Ferroxcube), essential for switching. Such switching action is obtained by magnetic circuit saturation or by base-current control from a coupled winding. The optimum operating frequency is between 1 and 5 kHz. For output powers of less than 5 and even 10 watts, one-transistor circuits are satisfactory; for higher powers, 2-transistor push-pulls are preferred. The push-pulls also provide a comparatively low output impedance.

The 40w 24-to-100v converter in Fig. 8-28 uses two push-pull connected transistors, with the bases also push-pull fed by a symmetrically coupled winding. The secondary voltage is rectified by a bridge rectifier. There are countless converter circuits more or less similar. For higher powers, external-control rather than auto-excited converters are used, and for inverters requiring frequency regulation, a stable, and possibly synchronizable driving oscillator is mandatory.

Since inverters and converters are switching-mode devices, thyristors will outdo transistors, especially for high powers (kilowatts of output). But even in 100w devices, thyristors provide efficiencies of some 90 percent.

As discussed in Chapter 3, if thyristors are DC-fed, as in inverters and converters, a means must be provided to de-energize them to end their operating cycle. In the 2-thyristor inverter in Fig. 8-29, this is obtained by a resonant LC circuit. Suppose SCR1 is conducting and SCR2 is blocked to begin with. The DC supply current then passes through the left half of the center-tapped output transformer (T1) secondary and, with T1 behaving as an autotransformer, twice the DC supply voltage appears at the anode of SCR2 and across capacitor C. SCR2 fires on the arrival of the next gate triggering pulse and momentarily brings the top terminal of switching inductance L to twice the supply voltage. This applies a strong inverse bias to SCR1 and de-energizes it. The latter, now blocked, can be fired by the next control pulse that will simultaneously block SCR2.

The shortcoming of this circuit is its vulnerability regarding purely inductive loads which place considerable voltage transients across the idling or slightly-loaded SCRs. Adding diodes D1 and D2 (as devised by McMurray and Bedford) provides a square output voltage and reduces

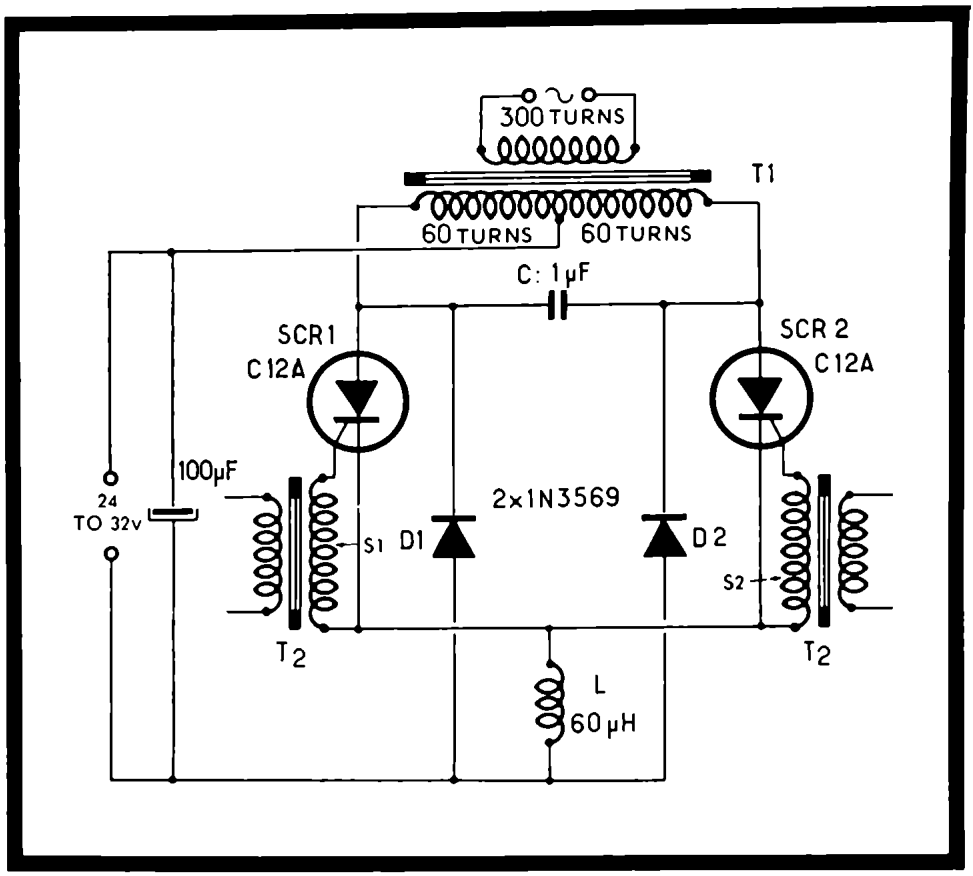
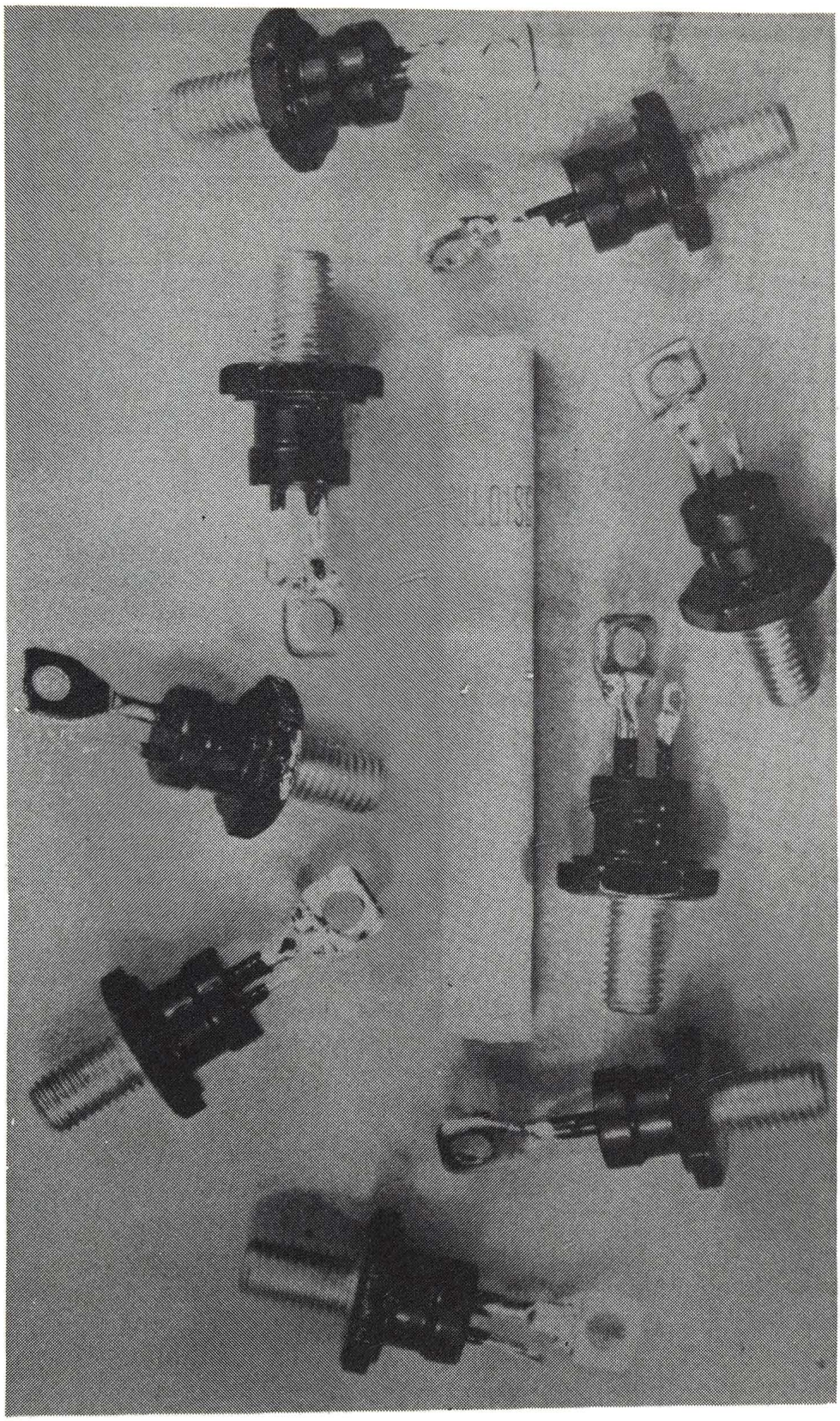


Fig. 8-29. This 2-thyristor inverter circuit is de-energized by resonant circuit LC.

voltage transients, thereby allowing for the use of (less expensive) SCRs with a lower PIV rating. With the discharge of C through L in the form of oscillations, if the anode of SCR2, for instance, runs negative, D2 begins conducting at the end of the SCR1 conduction cycle. The energy stored in L is then dissipated by D2, SCR2 and the circuit resistance.

To drive this inverter, General Electric points to the unijunction (UJT) trigger circuit in Fig. 8-30. Transistors T1 and T2 form a hybrid multivibrator applying an approximately square wave across the primary of transformer T2. Secondaries S1 and S2 inject trigger pulses into the SCR gates. This multivibrator is freely oscillating at about 100 Hz and is easily synchronized with a higher frequency by pulses produced by the UJT and injected by means of the common emitter resistor. A variable resistance allows for varying the sync frequency between 400 and 2,500 Hz and the multivibrator keeps in step. (If 60-Hz operation is desired, the multivibrator could oscillate at, say, 30 Hz, using a higher RC-product and a higher capacitance in UJT circuit.) Diodes D3 and D4 prevent the development of an excessive voltage across the base-



High-power thyristors rated at 35 amps.

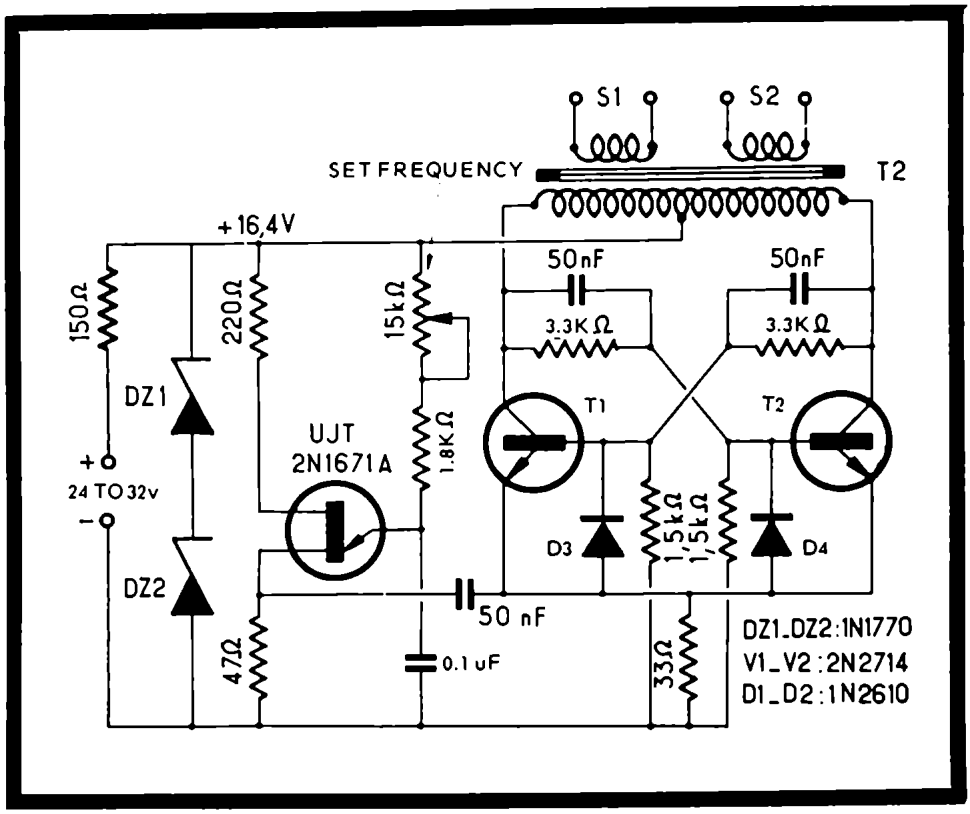


Fig. 8-30. Unijunction transistor inverter triggering circuit.

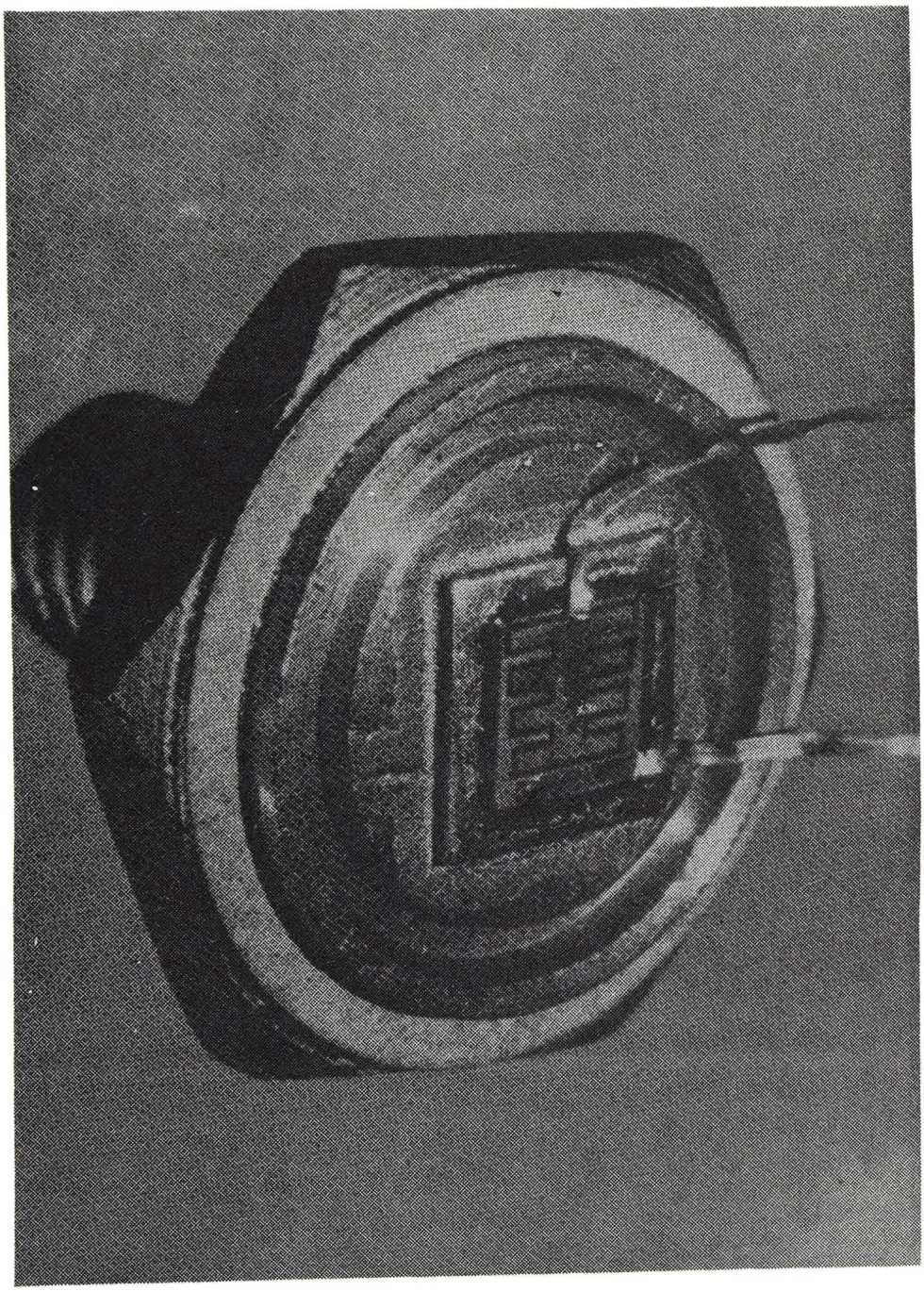
emitter junctions during the nonconduction period of the corresponding transistor. To reduce frequency drift, the driver stage supply voltage is regulated by zener diodes.

Three-phase inverters can be built using three thyristors (Fig. 8-31). A half-wave circuit is shown, and a full-wave 3-phase inverter using six thyristors is presented in Fig. 8-32. The corresponding trigger circuits are omitted to simplify the diagrams.

IGNITRONS

Ignitrons are cold-cathode controlled unidirectional devices such as thyratrons and thyristors, but they have much greater power handling capabilities. An ignitron comprises a water-cooled envelope (in spite of its high efficiency, the high power losses induce considerable heat dissipation) containing a pool of mercury that forms the cathode, a graphite anode and a conical tip made of boron carbide (the **ignitor**) extending slightly into the mercury pool (Fig. 8-33). Since a pair of ignitrons can handle up to 10,000 amperes, the cathode and anode connections obviously are rather sturdy.

Like a Thermionic tube, an ignitron can conduct only if its anode is positive with respect to its cathode. During its conductive half-cycle, it can be fired by connecting the ignitor to a positive voltage point such as the anode. Since there is considerable resistance between the boron carbide tip and the



100-watt power transistor with protective cover removed.

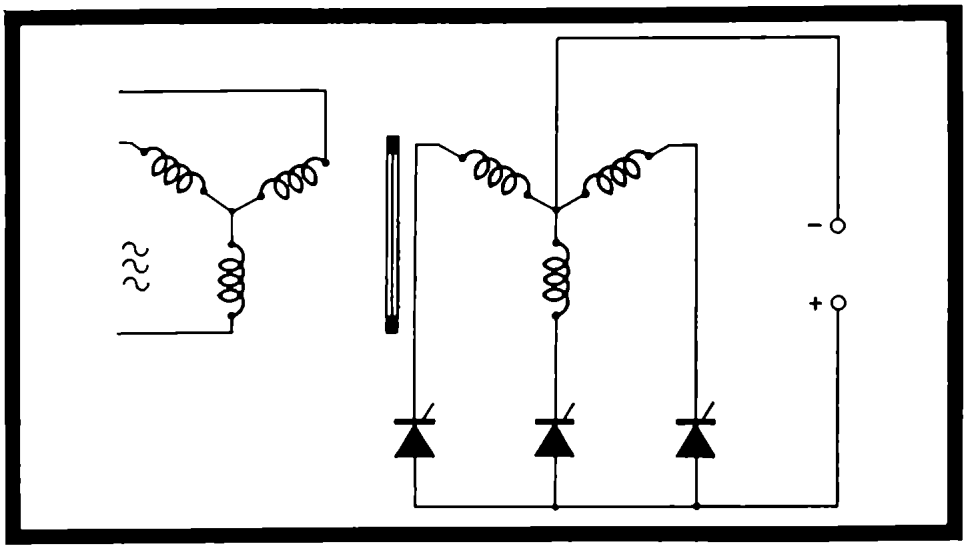


Fig. 8-31. Three-phase half-wave inverter using three thyristors.

nonwetting mercury pool, a very tiny arc builds up between them, releasing a sufficient quantity of electrons from the mercury to immediately ionize the cathode-anode space and start tube conduction. The discharge stops at the end of the half-cycle and is started again by the ignitor at the next positive half-cycle. For a very short time (some microseconds), the ignitor passes a rather heavy current (about 40A), but then the anode takes over, and the ignitor

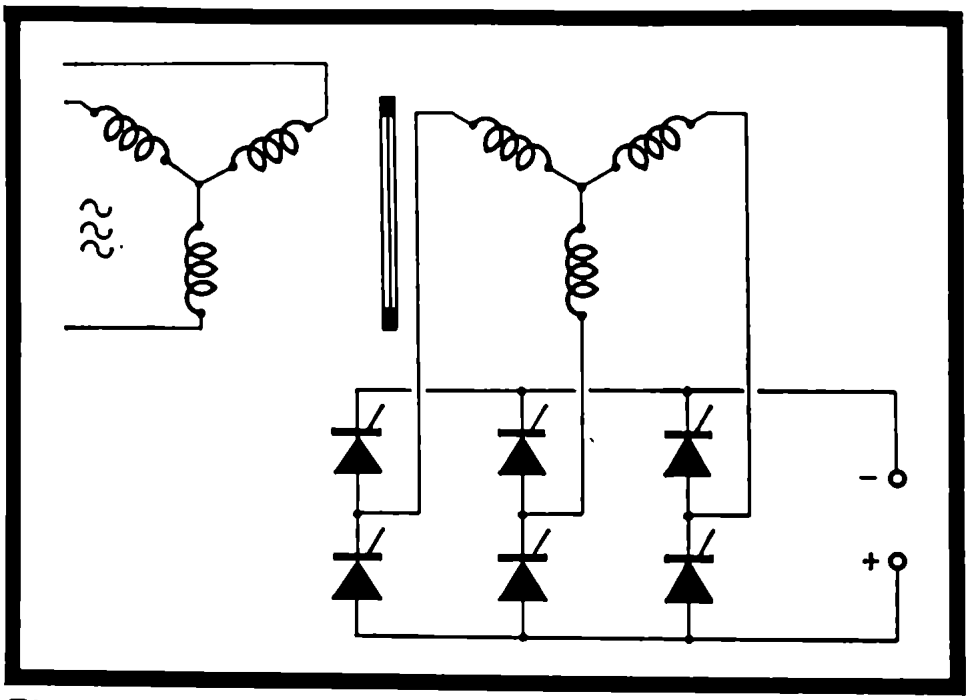


Fig. 8-32. The 3-phase full-wave inverter uses six thyristors.

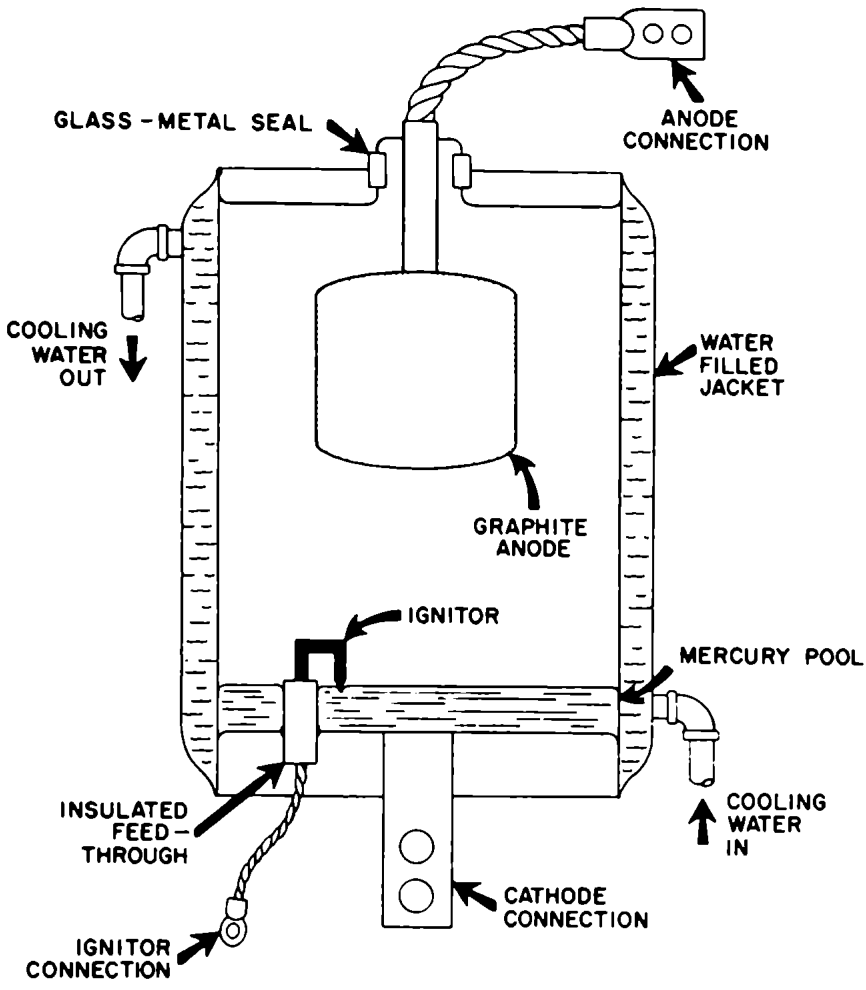


Fig. 8-33. Sketch showing construction of an ignitron.

current becomes negligible. Just as in a thyratron or thyristor, the discharge-initiating electrode (ignitor) loses control as soon as the device is conducting.

Because the high current evaporates a considerable amount of mercury, thereby increasing the vapor pressure, a protective device is provided to interrupt the power connection should the cooling water fail to circulate; a heat-sensitive bimetal strip will detect excessive heating. There are also thermostatically controlled "economizers" that shut off the water supply when the ignitron is cold (during idling periods).

Ignitron welders are described in Chapter 6. A single-ignitron half-wave rectifier or power switch is diagrammed in Fig. 8-34. As long as switch S is closed, the ignitor is powered through silicon diode D during the positive half-cycles and

A typical ignitron, the big brother of the thyratron.

triggers the main discharge. A unipolar voltage develops across load RL. During the negative half-cycle that follows, the diode prevents the current from uselessly passing through the ignitor and damaging it.

In the half-wave ignitron rectifier or switch in Fig. 8-35, the diode is replaced by a thyristor (Th) whose firing angle is

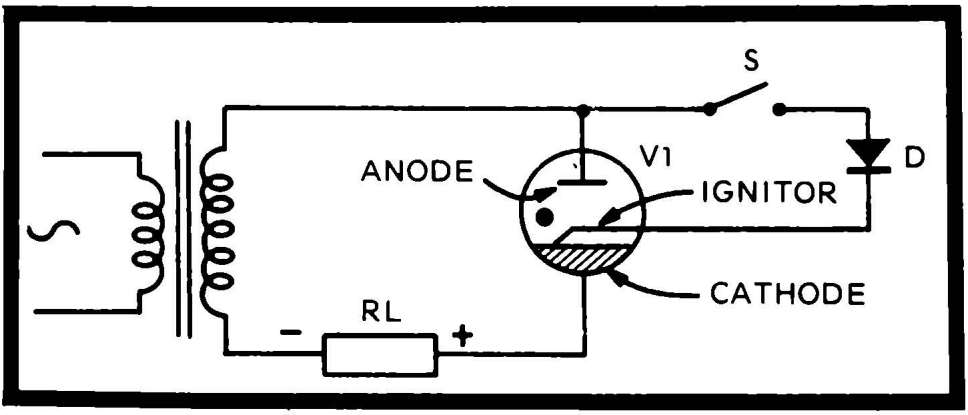


Fig. 8-34. Schematic of a single ignitron half-wave rectifier or power switch.

controlled by a phase bridge composed of the center-tapped secondary, resistor R and saturable reactor Ls used in place of a capacitor. A DC control current applied to Terminals XY varies the inductance of Ls, thereby providing horizontal control of the thyristor and ignitron Ig by the same token. During the negative half-cycle, Th is blocked and it protects the ignitor. While the preceding circuit is a mere on-off control, the addition of a thyristor allows for gradual power control.

A high-power half-wave 3-phase rectifier of switch is easily derived from the circuit in Fig. 8-36. There are three ignitrons tied to the wye-connected secondary, with three thyristors in the ignitor circuits. Because of the "hot" thyristors, separate trigger transformers are required for the three phases. A common DC control winding simultaneously shifts the phase angle of the three thyristors and the firing angle by the same token.

Ignitron Power Control

Ignitrons are extensively used to deliver accurately timed preset amounts of power, especially to high-power resistance welders. In Chapter 6, a full-wave welding control using back-

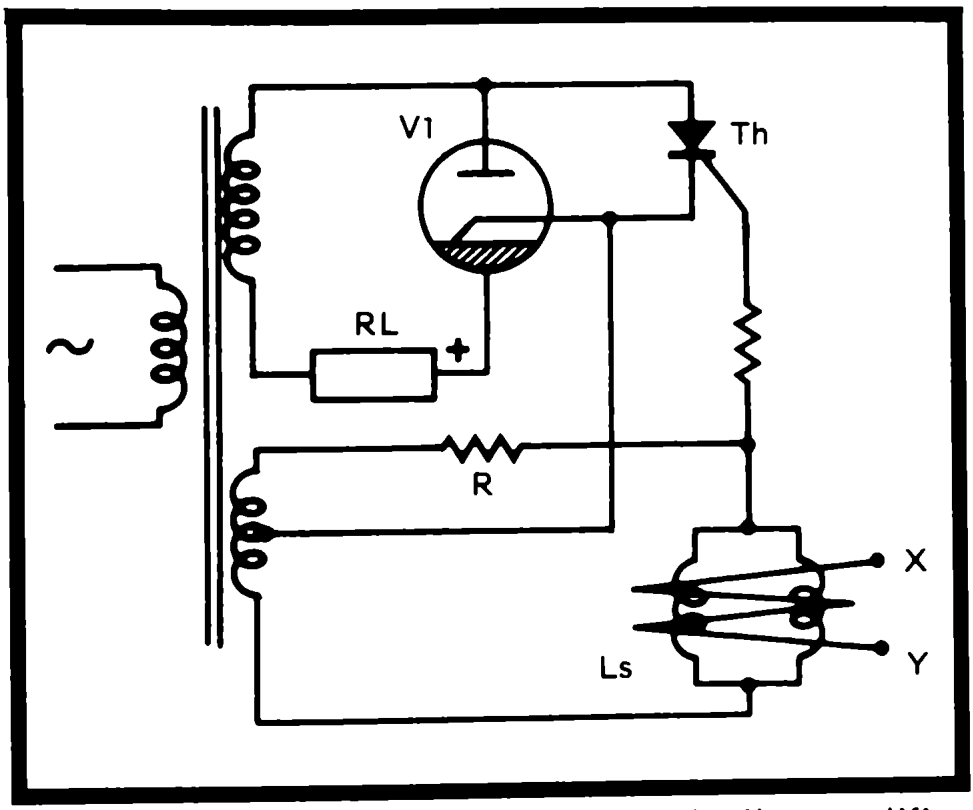


Fig. 8-35. Thyristor-controlled half-wave ignitron rectifier circuit.

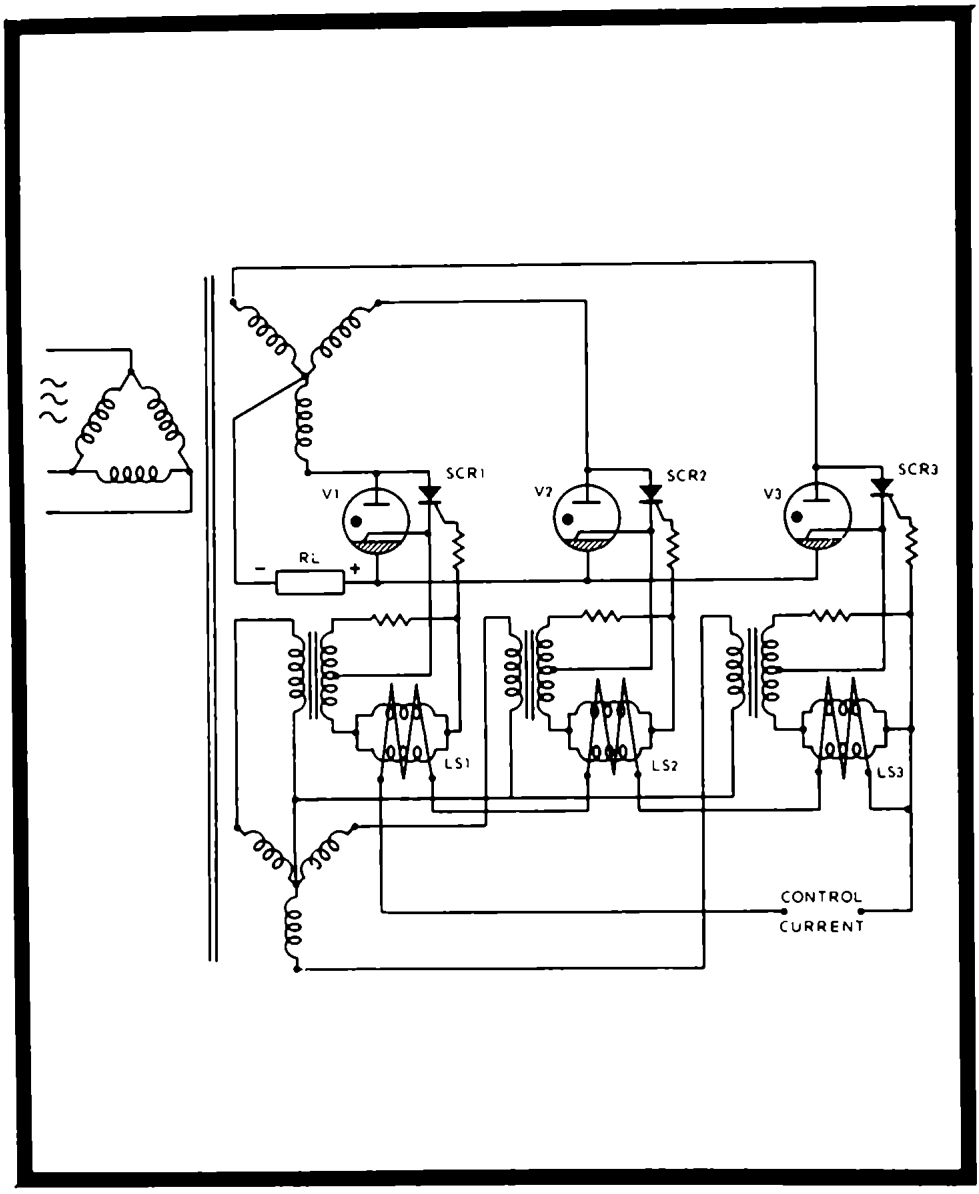


Fig. 8-36. Three-phase half-wave ignitron rectifier or switch circuit.

to-back connected ignitrons in series with the welding transformer primary are shown in Fig. 6-11. Thyristors control the ignitor current and the firing angle by the same token. The box labeled "welding control" contains ganged phase bridges similar to that in Fig. 8-35 for the two thyristors.

A pair of back-to-back connected ignitrons used, for instance, in a welding contactor can be made self-striking like the ignitron in Fig. 8-34. Consider the circuit in Fig. 8-37 for the half-cycle where power supply Terminal A is positive regarding B. Only V2 can conduct. The starting current is

passed from the ignitor to the anode of V2 through rectifiers V4 and V6, provided timing and water flow switches S1 and S2 are closed. Opening this loop disables the ignitor and stops conduction at the end of the half cycle. At the next half cycle, V2 remains extinguished, but V1 fires; its ignitor is energized through rectifiers V3 and V5. Such a contactor can be operated by a preset counter for a whole number of half cycles or full cycles. In contrast to continuously variable phase-shift control, the firing angle here is not variable: power is delivered for the full half cycle or not at all. For counter-operated welding controls, this is adequate.

Thyristors and Triacs of sufficient ratings to supply high-power welders are not yet readily available, but probably will be marketed sooner or later. For small- to medium-power applications, Triacs are attractive because a single device allows for full-wave operation, thus replacing back-to-back connected unidirectional switches. In the single-phase low-power welder in Fig. 8-38, a single Triac does all the power switching; it is triggered by a preset-counter controlled pulse generator. A 3-phase welder supply not only prevents power line unbalance during the weld phase, but also allows three

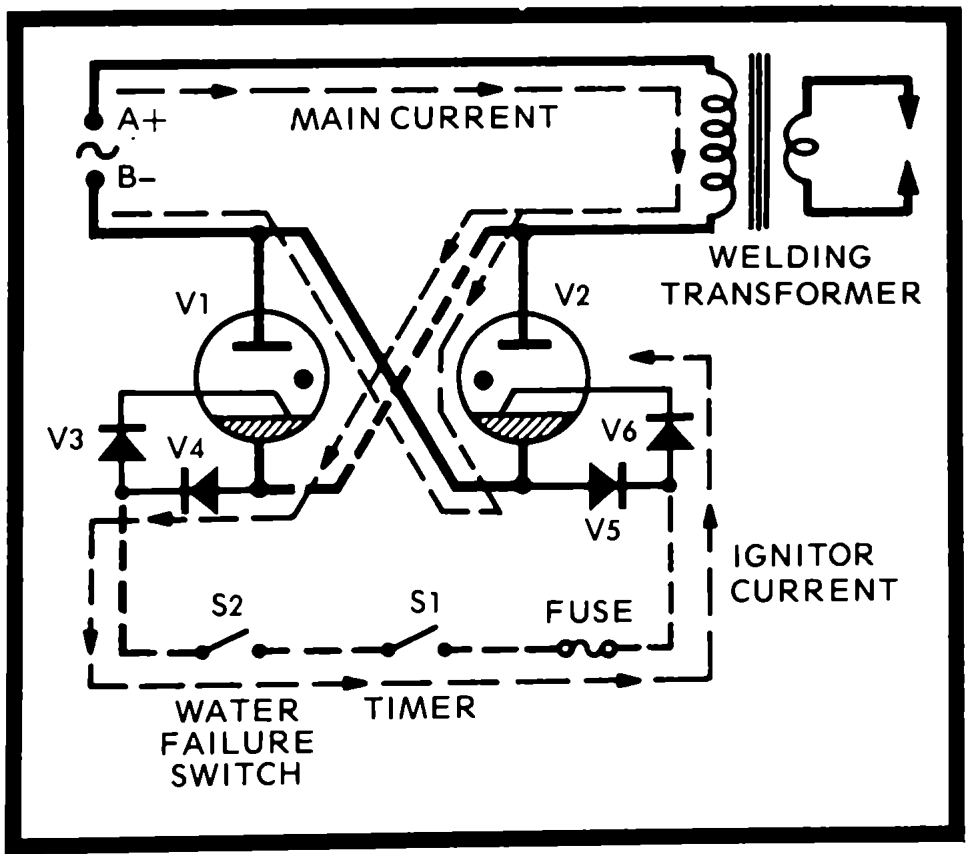


Fig. 8-37. Self-striking back-to-back connected ignitron pair used in a welder.

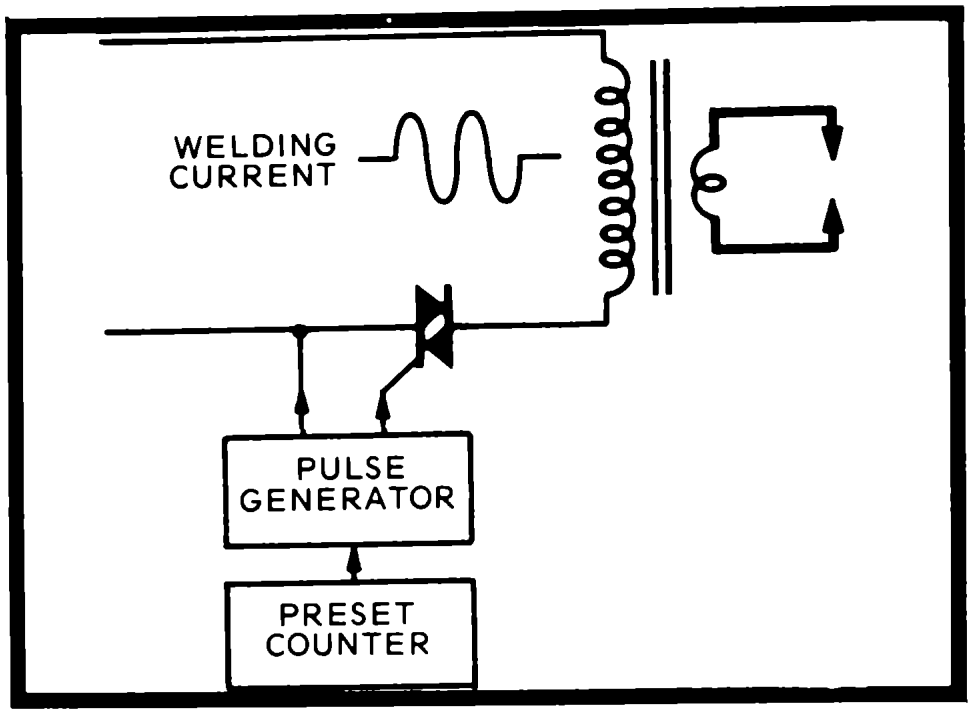


Fig. 8-38. Schematic of a single-phase Triac-controlled welding contactor.

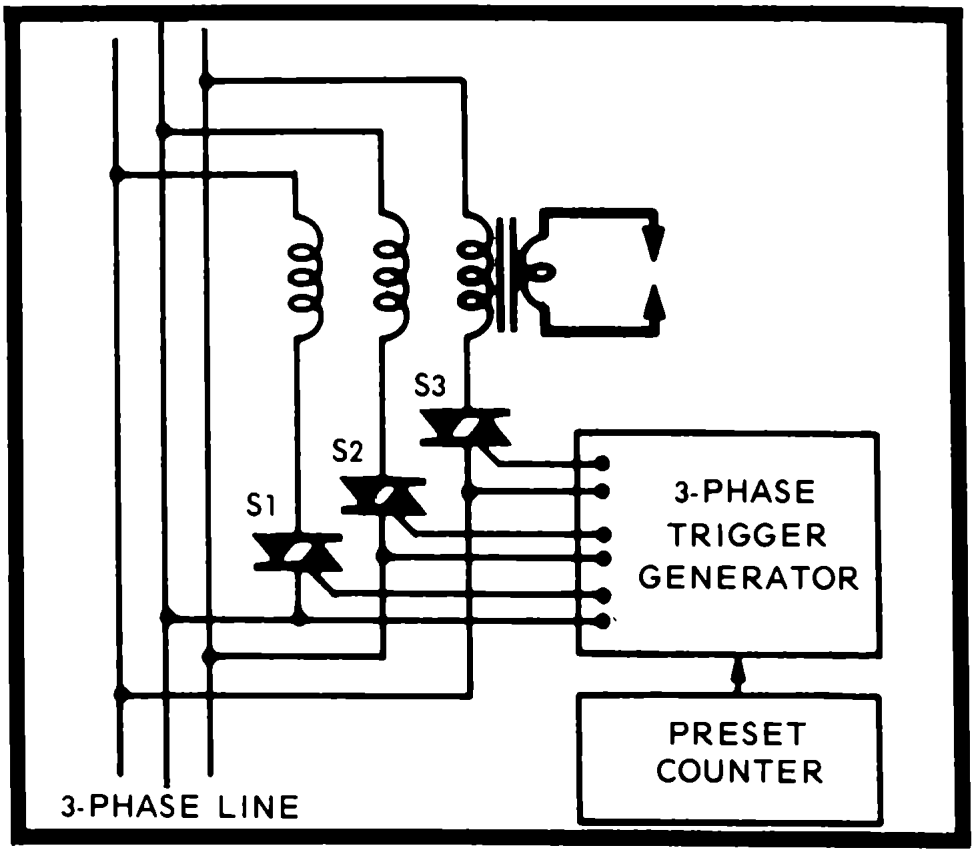


Fig. 8-39. Three-phase Triac welding contactor circuit using only three semiconductors.

times the rated power of one Triac to be delivered to the welder (Fig. 8-39). Each of the three switches must be provided an individual trigger generator and an insulating transformer, but the preset counter controls them all.

SATURABLE REACTORS

An iron-core inductor opposes an alternating current because of its inductive reactance $L\omega$ which is much higher than the ohmic resistance of the winding because the current flow must build up a magnetic flux in the core in opposing directions two times a cycle. Series connected with an AC-fed load RL (Fig. 8-40), such an inductor (L1) will introduce a voltage drop VL, reducing the voltage VR across RL. Since the ohmic resistance of L1 is generally negligible in comparison to the reactance of $L1\omega$, little loss appears as heat. Because of the different phase angles, VL and VR compose vectorially, and their sum exceeds input voltage Vi.

If a variable DC current is passed through a second winding (L2) wound on the same core, the magnetic laminations are more or less saturated. During half cycles of one-polarity, the alternating current traversing L1 does not have to magnetize the core and, therefore, passes more easily, considerably reducing the apparent inductance of the series winding; this increases the power delivered to RL. In a well

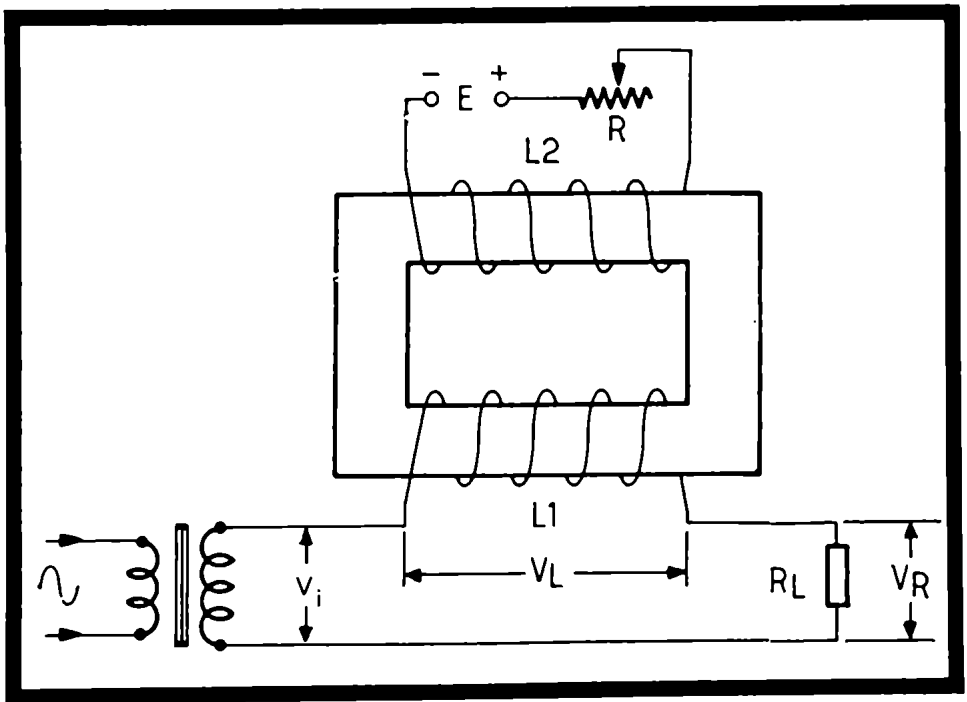


Fig. 8-40. Schematic illustrating the operating principle of a saturable reactor.

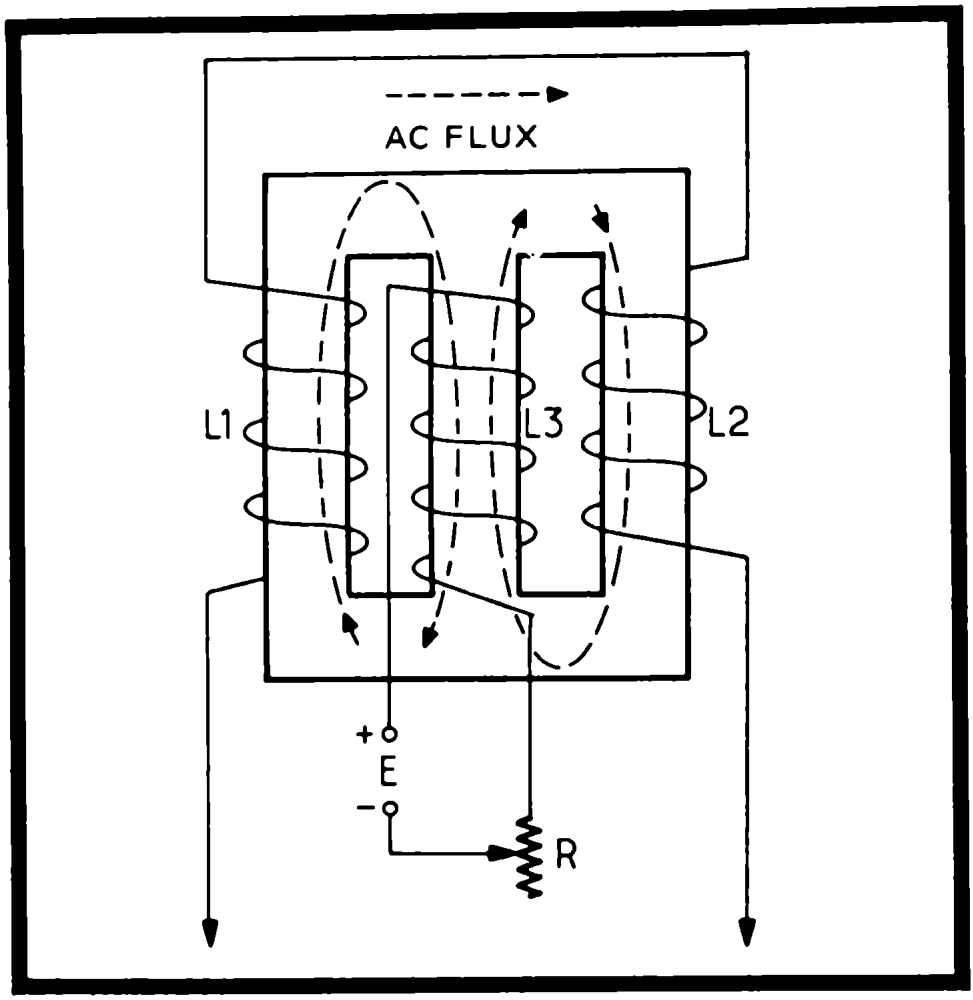


Fig. 8-41. In a 3-leg reactor, the AC flux cancels in the control leg.

designed saturable reactor, an appreciable load current can be varied within wide limits by a much smaller control current: this, obviously, is **current amplification.**

Ordinary transformer laminations are difficult to saturate, as indicated by the smoothly rounded hysteresis loop. For saturable reactor construction, special grain-oriented nickel-iron or silicon-iron laminations featuring a rather square hysteresis loop, such as shown in Fig. 5-37, are used. With such a core, only a few ampere-turns are necessary to run into saturation, provided there are no air gaps.

The simple arrangement in Fig. 8-40 is inconvenient to use, for the device is nothing other than a transformer. The AC voltage used to impede control action is induced into L2. This shortcoming is eliminated with the double-E core in Fig. 8-41. The load winding is split into two halves, L1 and L2, supported by the other legs. Depending on the impedance required, L1 and L2 can be series as well as parallel connected. The central

leg bears control winding L3, or perhaps several control windings. While the DC (control) flux saturates the entire core, the AC fluxes set up by L1 and L2 now oppose and cancel in the central leg; therefore, no AC appears across L3. The same goes for the 4-legged core in Fig. 8-42, where the two central legs support the two halves of the load winding and the control winding encloses both legs.

Saturable reactors are useful in many low-loss power control applications. They are used as light dimmers for theater floodlights and spots where rheostat control results in prohibitive heat dissipation and power losses. In such applications, however, saturable reactors are being superseded by thyristors because of higher efficiency, reduced bulk, and possibly a more attractive price.

Saturable reactors also are serviceable for voltage and current regulation purposes. In the power supply in Fig. 8-43, the DC output current traverses the control winding of the

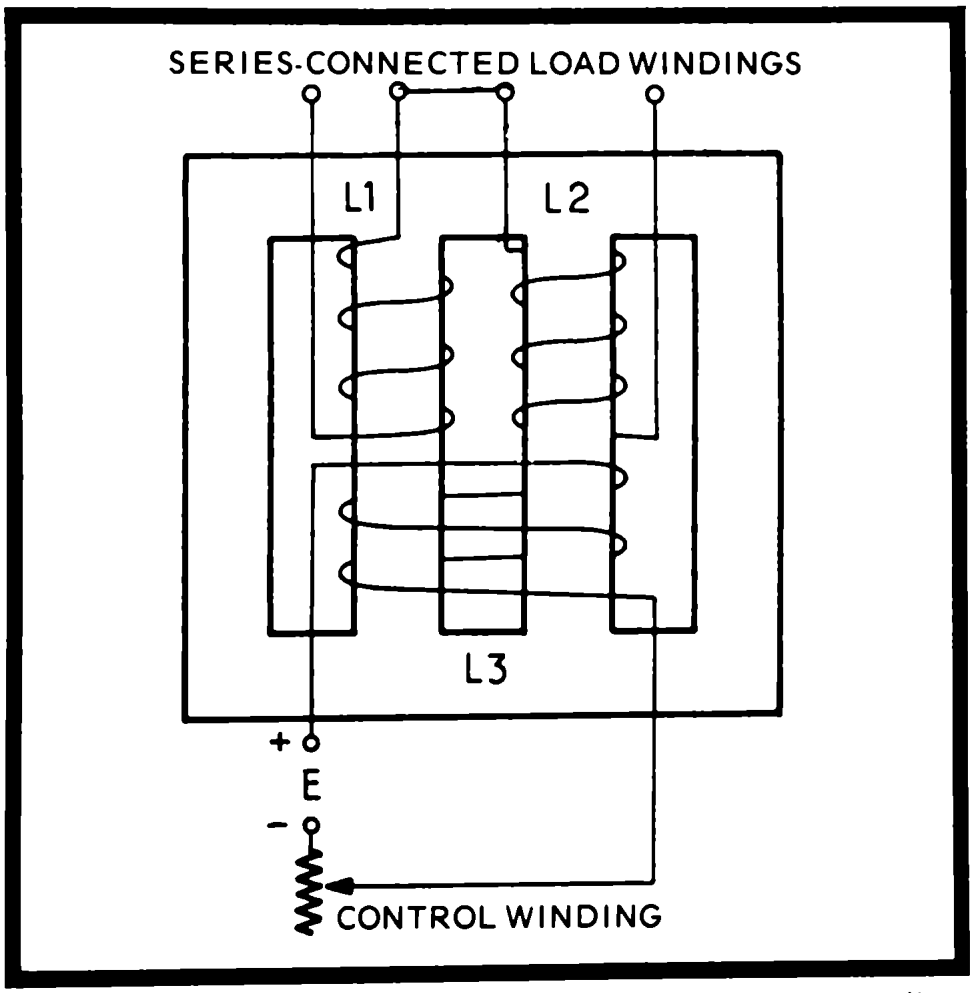


Fig. 8-42. The AC flux also cancels in the control winding of a 4-leg core.

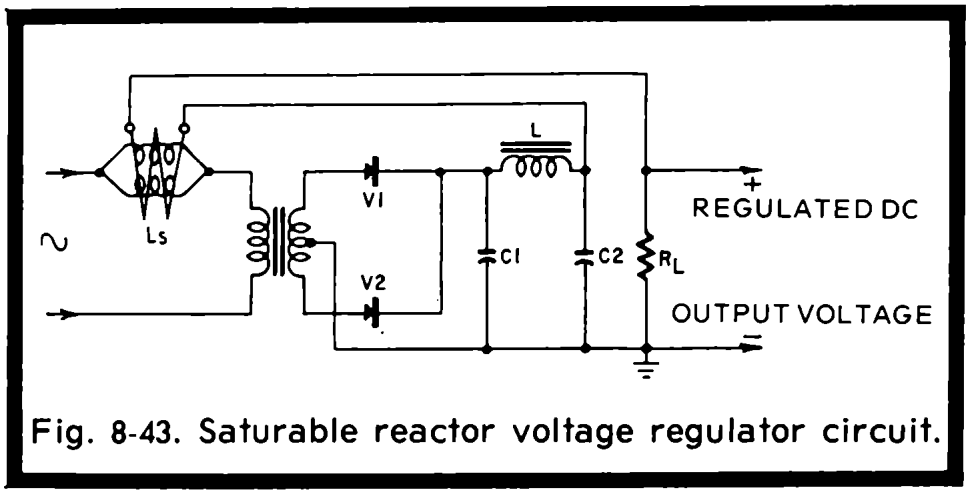


Fig. 8-43. Saturable reactor voltage regulator circuit.

saturable reactor which is series connected with the transformer primary. If the current drain increases, the DC voltage decrease can be overcome by an increase of the AC voltage across the primary; this is a constant-voltage or constant-load regulator. An AC voltage regulator using the same principle is diagrammed in Fig. 8-44. The transformer secondary voltage is rectified to energize the control winding of a saturable reactor in series with the transformer primary to compensate for an output voltage decrease by increasing the primary voltage. The efficiency of such a regulator can still be increased by skillful use of nonlinear resistors, such as varistors.

MAGNETIC AMPLIFIERS

Due to its amplifying properties, the saturable reactor already is a magnetic amplifying device, and it forms the

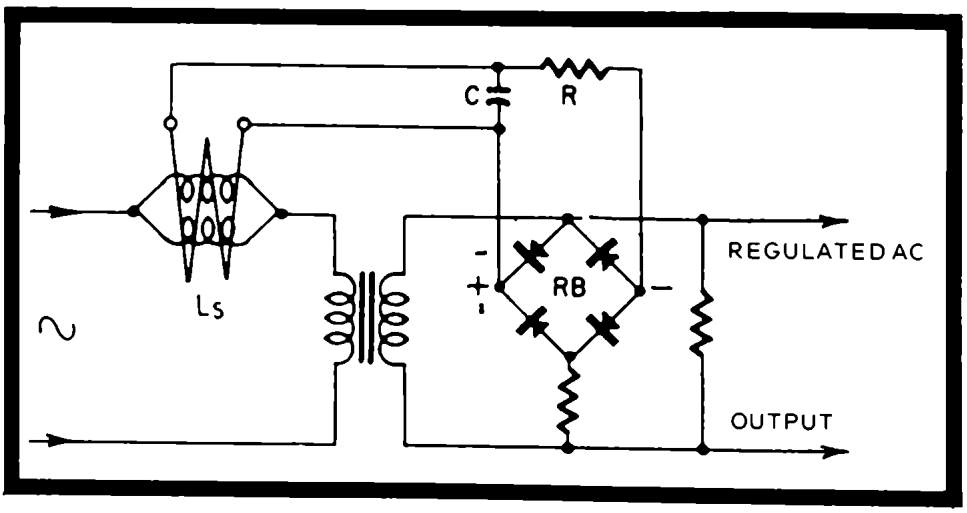


Fig. 8-44. AC voltage regulator using a saturable reactor.

principal part of a magnetic amplifier. In fact, it is not easy to find a neat boundary between these two devices. Some magnetic amplifiers are truly amplifiers, because a DC output current is controlled by a DC control current; but in many others, there is an AC output and the control current is still DC.

Since core saturation by the control current can only have an effect on the load current during the half cycles producing a magnetization of the same direction as the control current, the other half cycles are practically unaltered by the control current, resulting in a limited load-current variation. Added rectifier D (Fig. 8-45) passes only the half cycles that can be controlled. (It is the presence of at least one rectifier that distinguishes a magnetic amplifier from a saturable reactor.) This is a half-wave magnetic amplifier, and Fig. 8-46 shows (hatched) the load current variation during an AC cycle for an increasing control current (from top to bottom).

The magnetic amplifier in Fig. 8-47 is a full-wave device, for the two halves, L1 and L2, of the load winding are reversed regarding the AC supply and alternately pass current through rectifier D1 or D2, but always in a way to produce a flux of the same polarity as the control flux. As shown in the diagram in Fig. 8-48, for a control current increasing from top to bottom, the power delivered to the load is continuously variable from a very low level to full power. These amplifiers are self-

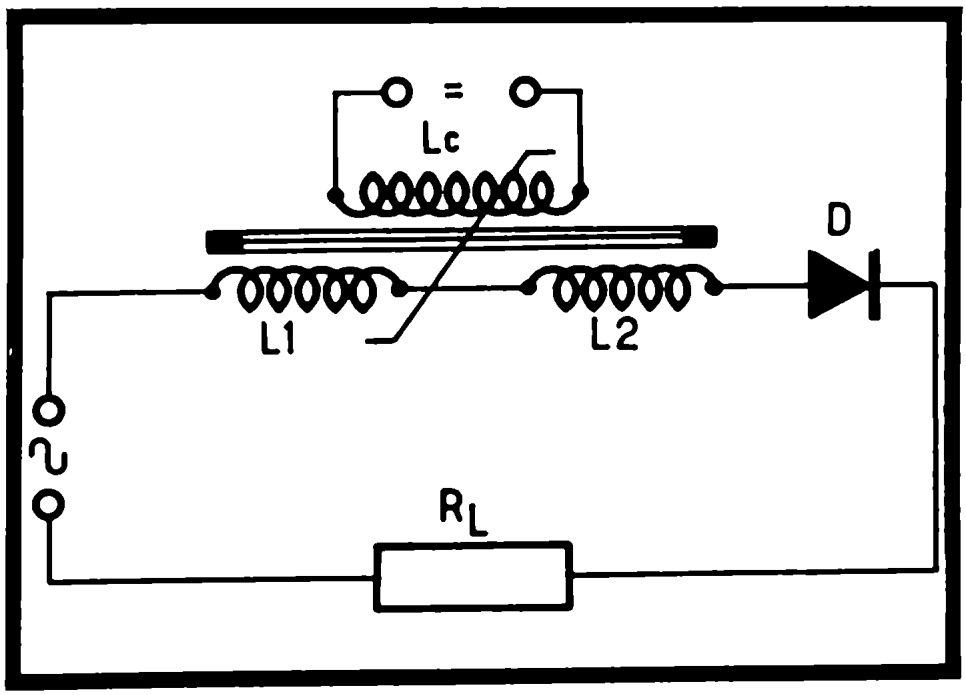


Fig. 8-45. Diagram of a self-saturating half-wave magnetic amplifier.

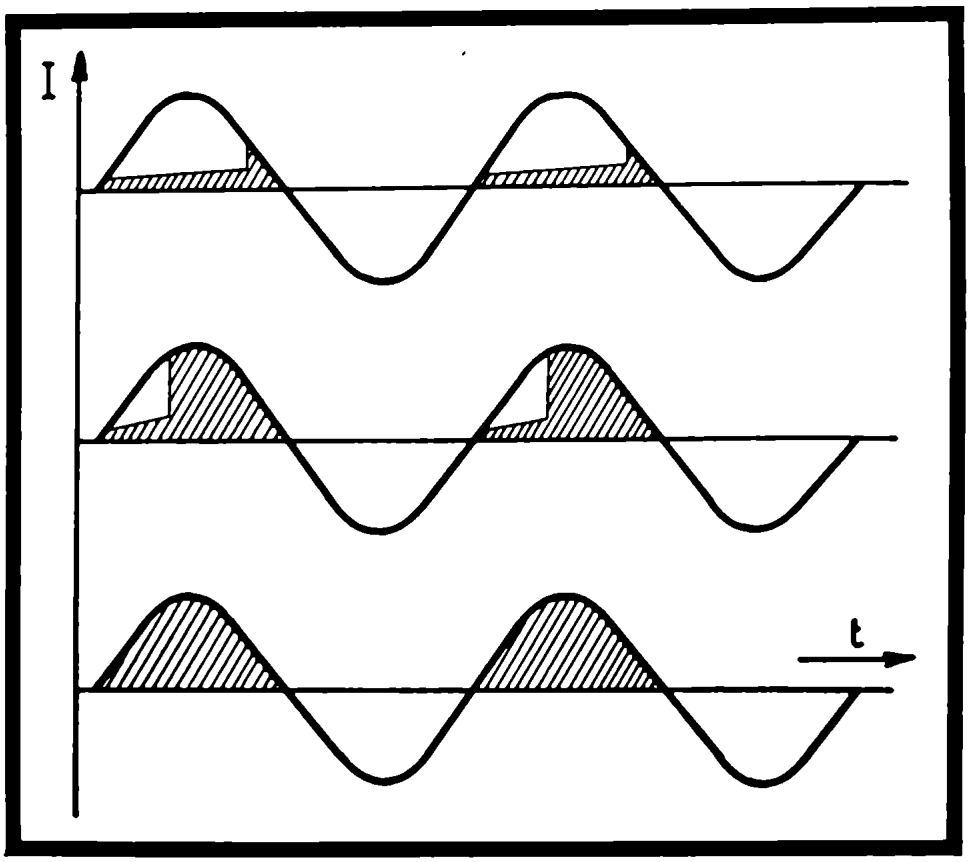


Fig. 8-46. These waveforms illustrate the load current variation in a half-wave magnetic amplifier.

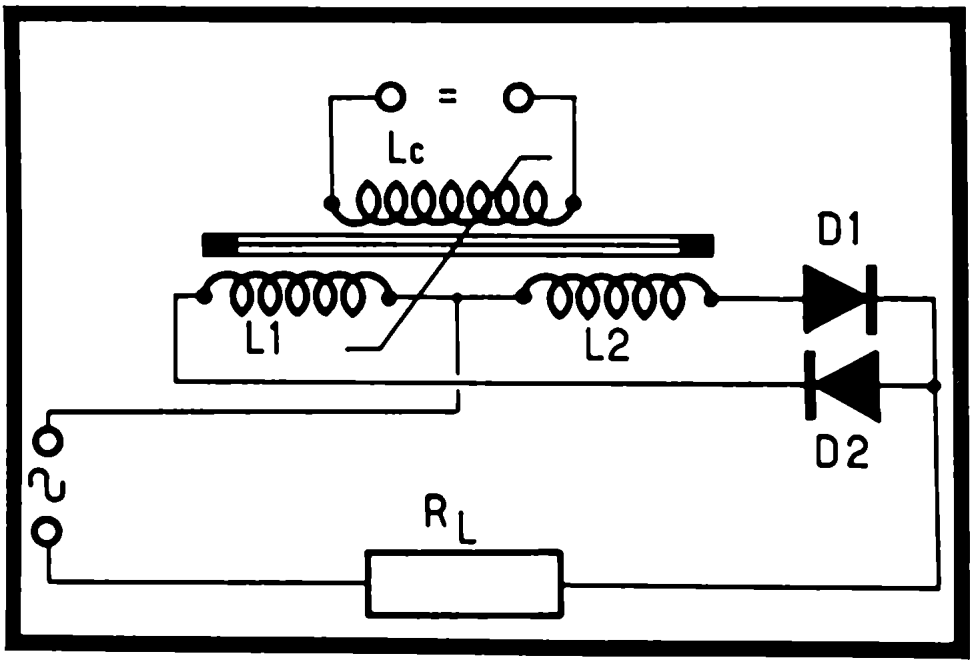


Fig. 8-47. Schematic of a self-saturating full-wave magnetic amplifier.

saturated by the half-cycle load current passing in the load windings, a feature saturable reactors are lacking.

The magnetic amplifier in Fig. 8-49 also is a full-wave self-saturating type, but it uses four open-bridge connected rectifiers. Since load RL is connected across the DC diagonal, it is energized by a unidirectional current. Hence, this type of magnetic amplifier is serviceable for DC motor control.

Just like its electronic counterpart, a magnetic amplifier can be subjected to a feedback that considerably modifies its characteristics and, especially, its gain. Positive feedback is obtained by applying a rectified part of the load current to a feedback winding (Lr) so that the feedback flux adds to the control flux. This is done in the **external-feedback** self-saturating magnetic amplifier in Fig. 8-50. The load current passes through a rectifier bridge whose DC diagonal applies a unidirectional current to feedback winding Lr, thereby reinforcing the action of control winding Lc. This results in current gains of 200 to 300 (against 10 with no feedback) and power gains of 104 to 105 (against about 40)—an amazing increase. In the circuit in Fig. 8-50, RL is AC-fed. To obtain a unidirectional load current, RL is series-connected with

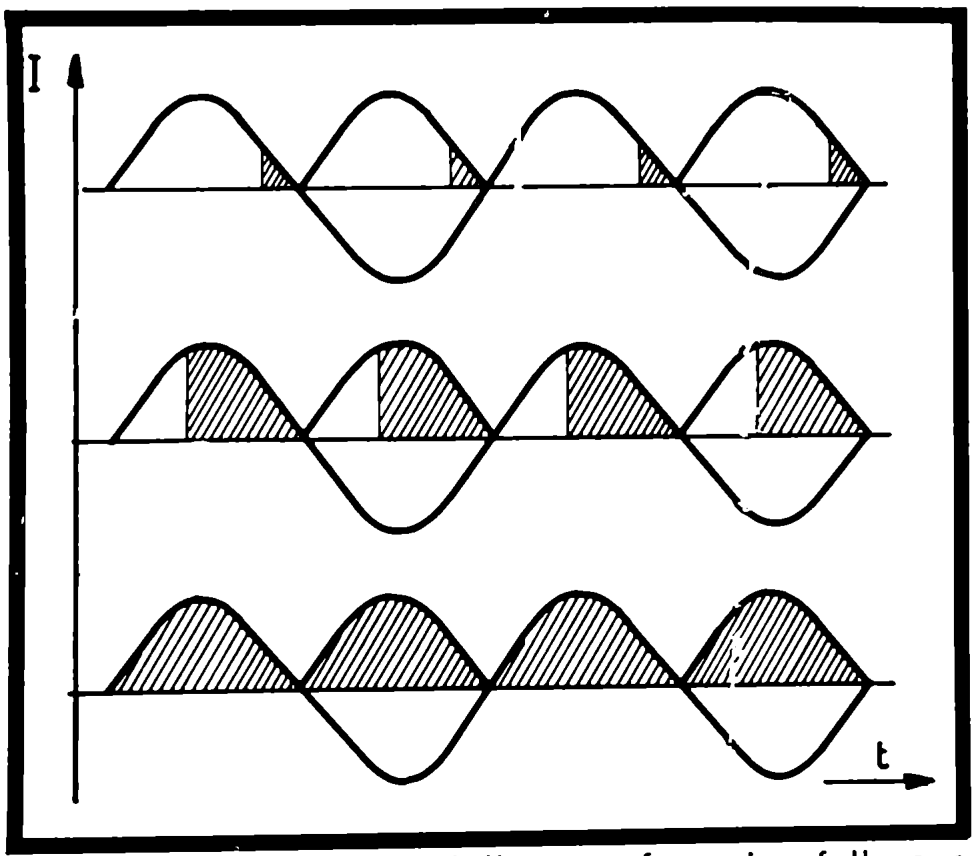


Fig. 8-48. Load current variation waveforms in a full-wave magnetic amplifier.

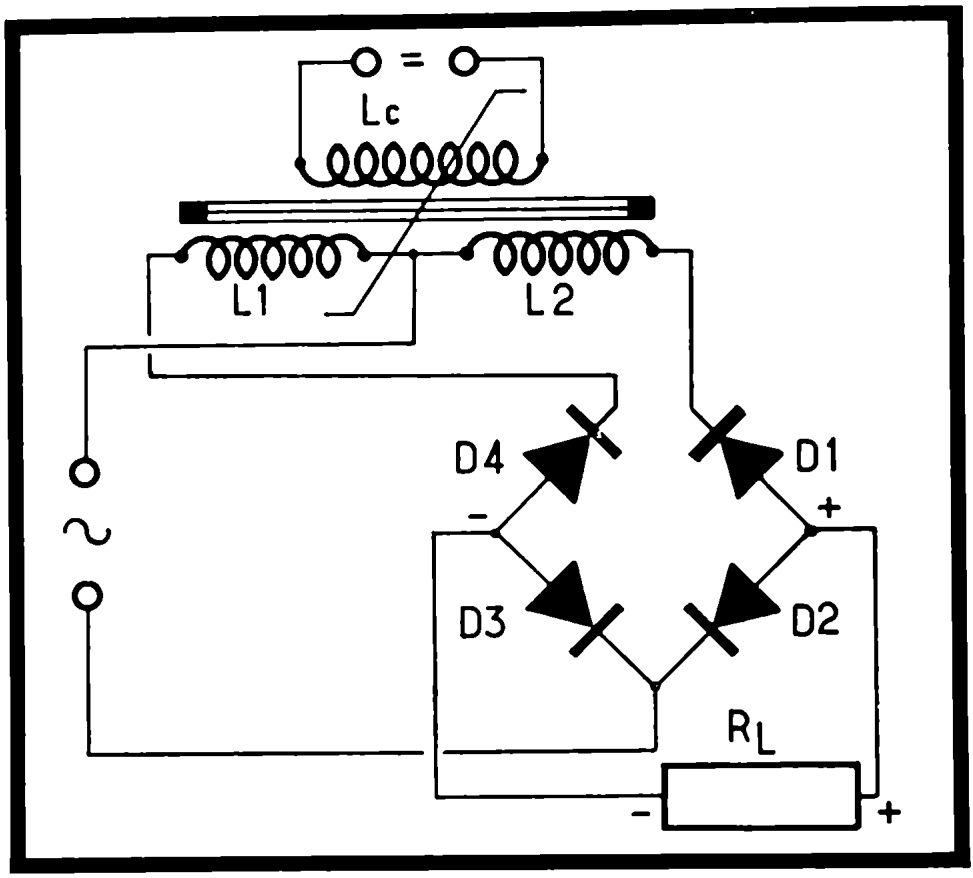


Fig. 8-49. Schematic of a full-wave self-saturating magnetic amplifier designed to deliver a unidirectional load current.

feedback winding Lr on the DC-side of the rectifier bridge (Fig. 8-51).

In the self-saturating amplifier in Fig. 8-47, an internal feedback is present due to the load currents in windings L1 and L2; diodes D1 and D2 rectify the currents in L1 and L2 according to the instantaneous input terminal polarity, a DC current, proportional to the mean alternating current value, passes through L1-D2 or L2-D1, and the windings are connected to reinforce the control current action. With this circuit, the load is AC-fed; however, it can be unidirectional-current operated by using additional rectifiers.

All these amplifiers are designed to vary the load current within wide limits; the current however, cannot be made zero, for the load winding impedance cannot be made infinite. And with a unidirectional (pseudo-DC) load current, it is not possible to reverse its polarity according to the direction of the control current.

A magnetic amplifier featuring an output voltage proportional to the control current and in which the polarity is

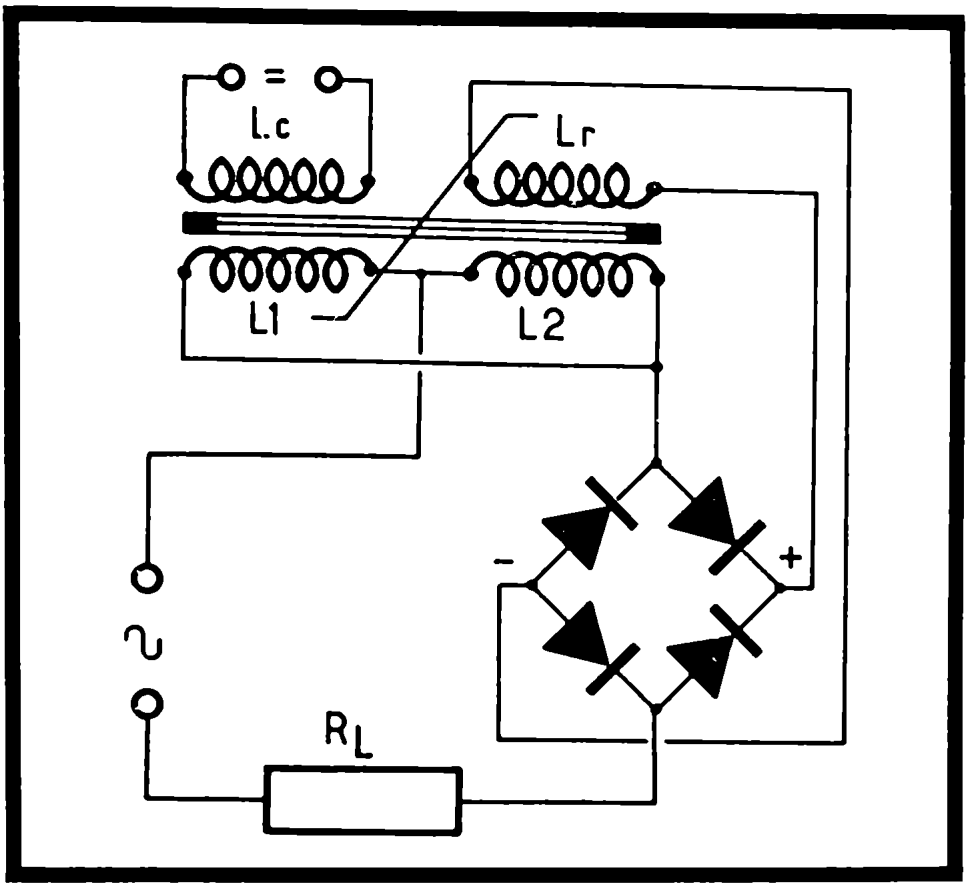


Fig. 8-50. Diagram of an external feedback self-saturating magnetic amplifier.

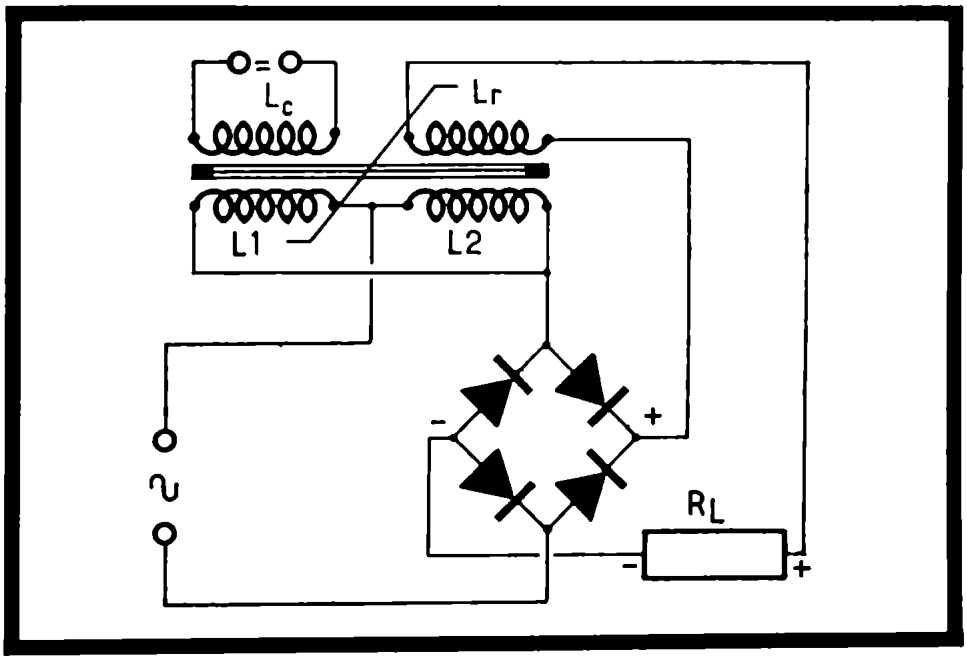


Fig. 8-51. External feedback magnetic amplifier circuit with a unidirectional output.

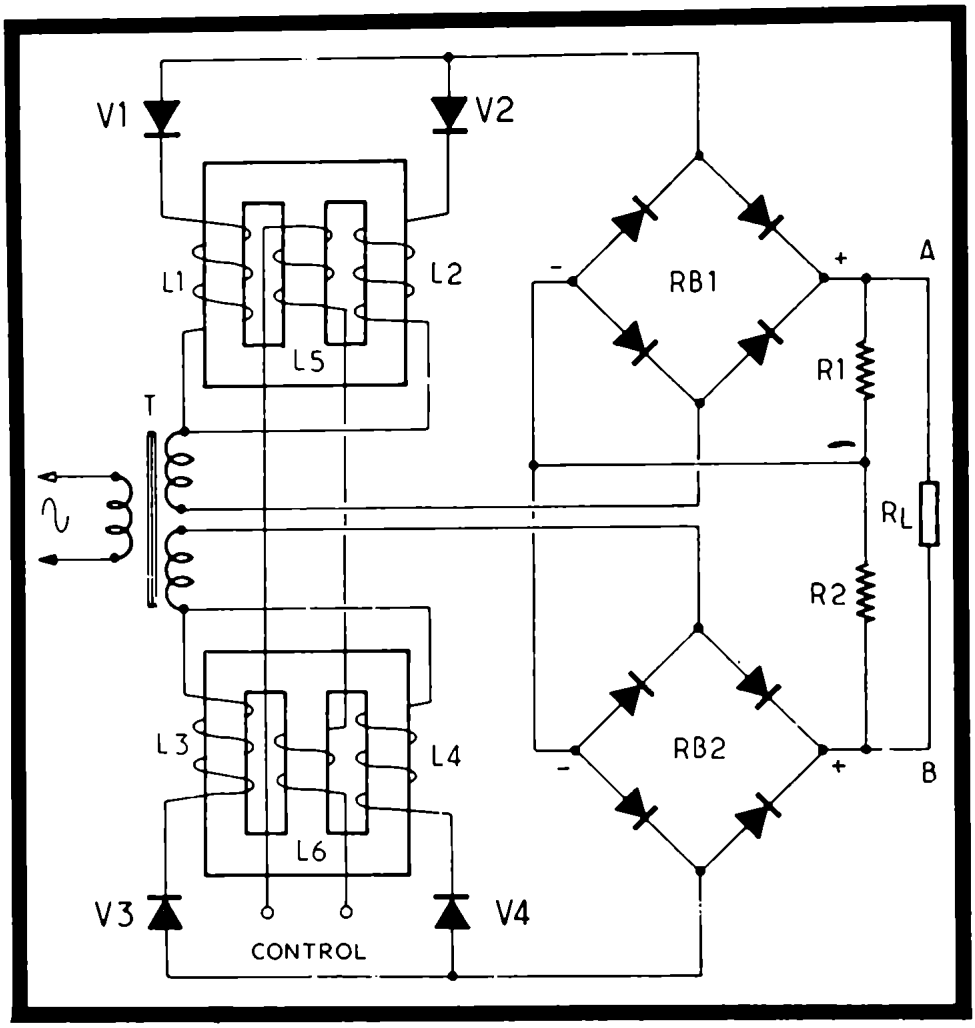


Fig. 8-52. Tandem magnetic amplifier circuit featuring a reversible polarity output.

restored can be achieved by coupling two similar magnetic amplifiers (actually, of the self-saturating type) in a way to oppose the output voltages rectified by the rectifier bridges (Fig. 8-52). With no control current passing in series-connected control windings L5 and L6, equal voltages develop across output resistors R1 and R2, but because of the opposing polarities, Points A and B are at the same potential, and there is no voltage across load RL. L5 and L6 are connected in a way to increase the saturation of one core for a given input while decreasing that of the other core; this will vary the voltages across R1 and R2 in opposite directions, so that the overall output voltage across RL can be positive, negative or zero.

A magnetic amplifier is a static device not subjected to wear; it can be potted and sealed to resist adverse environment and its life expectancy is rather high. Its power ranges from a couple of milliwatts to many kilowatts. Several

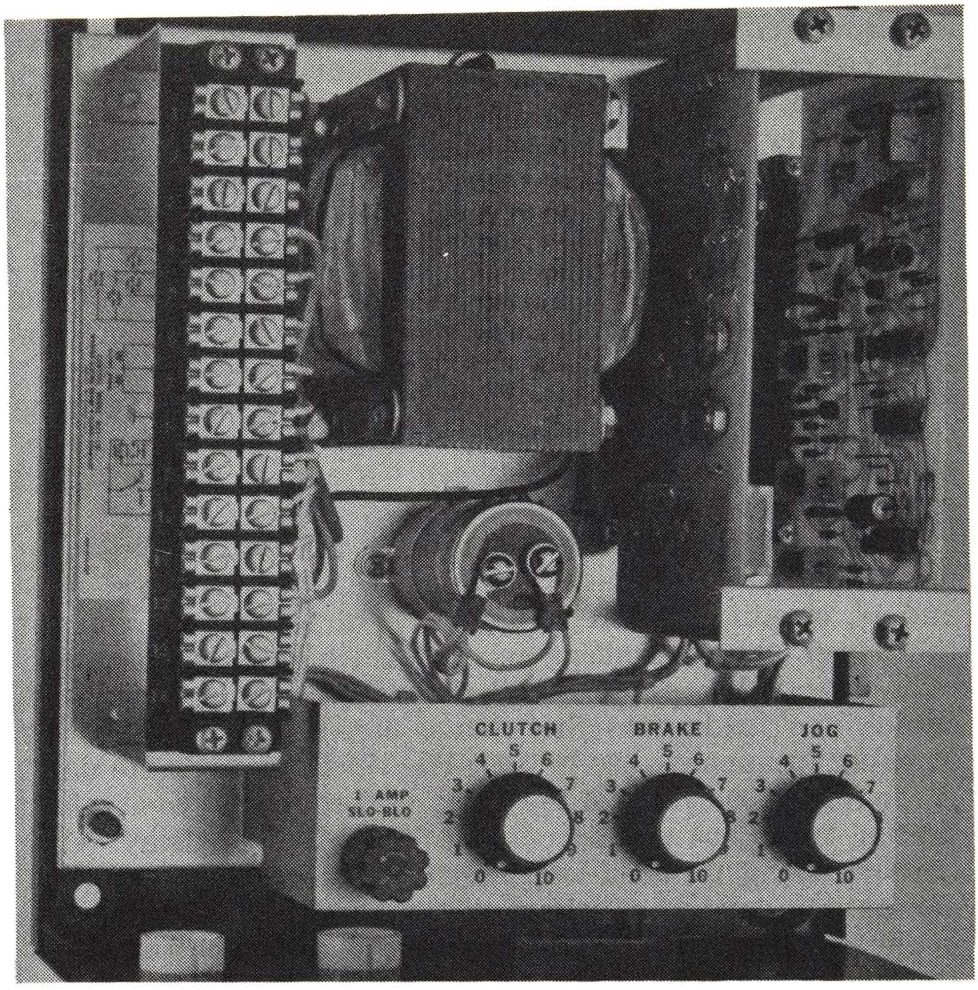
amplifiers can be cascaded to obtain a high power gain, and such cascaded stages are actually preferred to a single high-gain stage. Its bulk is approximately three times that of a transformer of similar power rating, and it is inversely proportional to the supply frequency (generally comprised between 50 and 2,000 Hz). Control input impedance is low, from 0.1 to 3,000 ohms according to the winding(s), and a magnetic amplifier matches fairly well low-impedance transducers such as thermocouples, photoconductors and strain gages. Output impedance generally ranges between 10 and 2,000 ohms at 50 or 60 Hz and somewhat more at higher frequencies. Drift is practically negligible, and stability and linearity are somewhere between 0.5 and 10 percent of the maximum load current.

The shortcoming of a magnetic amplifier is its comparatively slow response. The time constant depends on DC and AC circuit characteristics and is directly proportional to the power gain (resulting in sluggish high-gain amplifiers) and inversely proportional to the supply frequency, sometimes leading to the choice of a comparatively high working frequency (500 to 2,000 Hz). Practically, time delays of from 2 to 100 cycles of the working frequency are obtained; that is, from 0.04 to 2 seconds at 50 Hz. Such a delay can be embarrassing in closed-loop control systems. (Ramey devised a magnetic amplifier whose response time equals a single cycle of the supply frequency—at the price, however, of a smaller power gain.)

When cascading two amplifier stages, the total time constant approximately equals the sum of the individual time constants, while the overall gain is the product of the partial gains. This explains that, for a given gain, the response time is reduced by cascading several stages instead of using a single high-gain amplifier.

DC MOTOR CONTROL

Because of their inherent flexibility and widely variable speed and torque characteristics, DC motors are overhelmingly used in industrial drives. According to the various connections between the **armature** (rotor) and **field** (stator), there are series, shunt and separate-field DC motors. All such motors can be reversed by inverting the armature or field connections. While shunt and separate-field motors will run only on DC, series motors can be DC as well as AC-fed, and, therefore, are sometimes referred to as "universal" motors, especially fractional-horsepower types. Tiny intrument and toy motors generally have a field set up by a



Electric clutch-break control designed for soft starting and stopping of high-friction loads and for positioning in small increments. An adjustable pulse train, lasting 5-seconds rather than a varying DC voltage, controls the rate at which electric clutches or brakes develop full torque. Acceleration, deceleration, and jogging are all individually adjustable. A brake or clutch can be controlled independently, or two brakes or two clutches can be controlled, one unit on at a time. Originally developed for textile machinery, the unit can be used in conveyor drives, materials handling equipment, and machine tools. Made by the Warner Electric Brake & Clutch Co.

permanent magnet; such motors, generally battery-powered, will run only on DC.

Series motors feature a high starting torque (essential for railway traction), but their speed control is poor. For low-power servo controls, the split-field DC series motor is attractive because it readily leads to a very simple differential control.

The speed of a shunt or separate-field motor increases with an increasing armature current and-or with a decreasing field. Up to a so-called base speed, the best speed control is obtained by armature voltage control; for still higher speeds, the field current is reduced to increase the speed. In electronic motor controls, the armature and field are generally supplied by separate rectifiers, conventional or controlled, so that the so-called shunt motor indeed becomes a separate-field motor. Since the field power is only a fraction of the armature power, the field rectifier usually has a much smaller power rating.

A thyristor half-wave series motor control is schematically represented in Fig. 8-53. The motor speed can be controlled by varying the conduction angle of thyristor SCR, inserted between the field and armature. The rotating armature generates a counter emf which reduces the SCR conduction angle—adequate polarity provided—and thereby tends to stabilize the speed. During the negative half cycle, capacitor C1 partly discharges through speed-setting

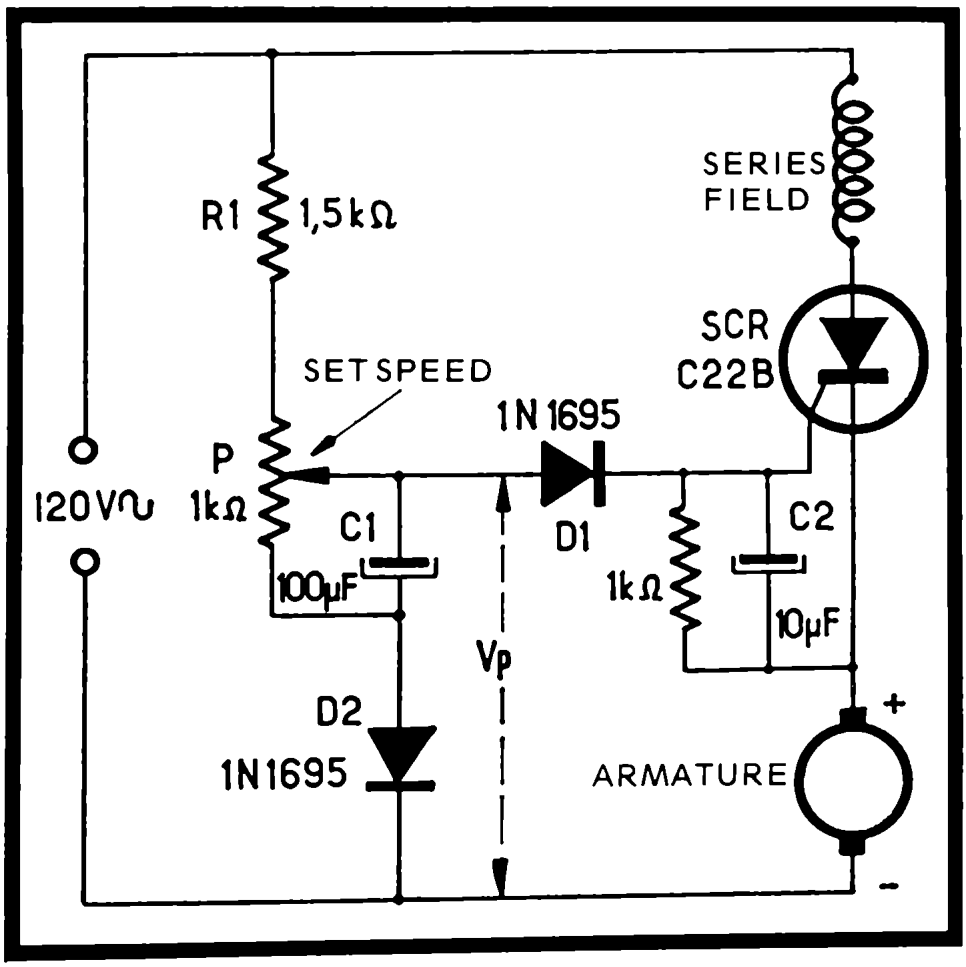


Fig. 8-53. Thyristor half-wave series motor control circuit.

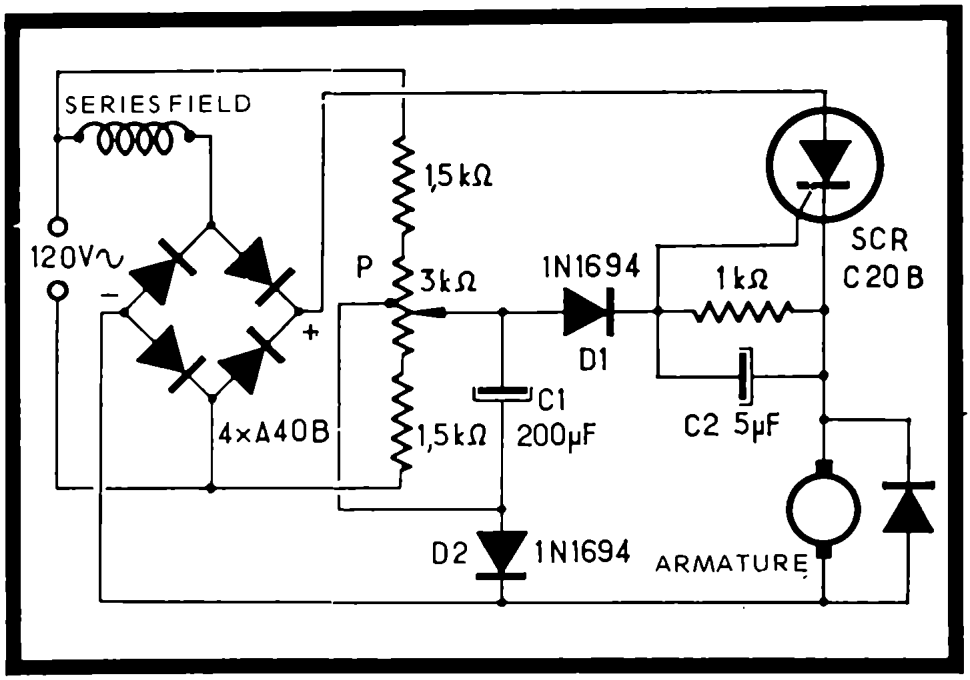


Fig. 8-54. Schematic of a reversible thyristor half-wave universal motor control.

potentiometer P; it recharges through R1 during positive half cycles. If the motor is stalling, there is no cemf and the SCR gate passes a current as soon as speed control voltage Vp exceeds the voltage drop of diode D1 and the SCR gate. Thus, the thyristor is triggered at the beginning of the half cycle, providing full power for motor starting. The running motor develops a cemf (the upper brush is positive) that opposes the triggering current supplied by C1 as long as it exceeds Vp. This retards the firing angle and allows the motor to cease accelerating. Once the motor has reached the preset operating speed, the armature voltage provides automatic speed regulation. (This motor can be reversed by inverting the armature or field connections; however, the regulating action will not only be lost but even thwarted.)

The thyristor half-wave universal motor control in Fig. 8-54 is a reversing type, for the field is connected on the AC-side of the rectifier bridge whose DC supplies the SCR and armature with positive half cycles (no rectifier output capacitor being provided). Center-tapped potentiometer P not only sets the running speed by varying the SCR conduction angle, but also the direction of rotation, according to firing taking place during the "true" positive half-cycles or the negative half-cycles after passing through the rectifier. Speed regulation here is operative for but one direction of rotation. Setting the slider to the center tap stops the motor. While the performance

achieved by these circuits is adequate for most universal motor applications, a higher degree of speed regulation can be achieved by more sophisticated circuitry.

A reversible half-wave speed control for separate or permanent-magnet field motors is schematically represented in Fig. 8-55. The armature is now series-connected on the AC line with two half-wave circuits, SCR1-D1 and SCR2-D2 back-to-back, and thus is fed by positive or negative half cycles according to whether SCR1 or SCR2 is conductive; this fixes the direction of rotation, and the field polarity remains unchanged. The nonconducting thyristor is shorted by a diode connected with the opposite polarity needed to conduct.

A unijunction transistor (UJT), controlled by reference voltage bridge R1, R2, D5 and D6, provides trigger pulses for

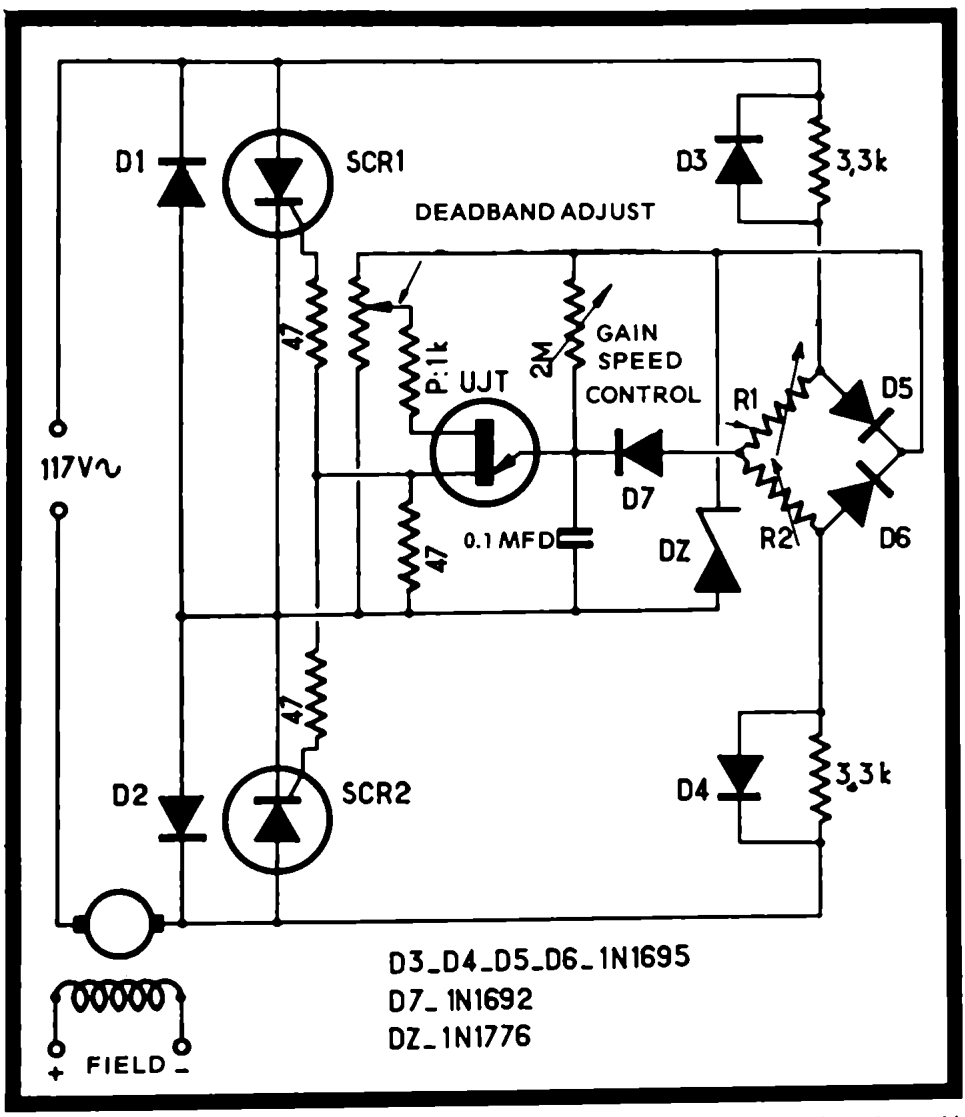


Fig. 8-55. Reversible half-wave speed control circuit designed for separate-field motors.

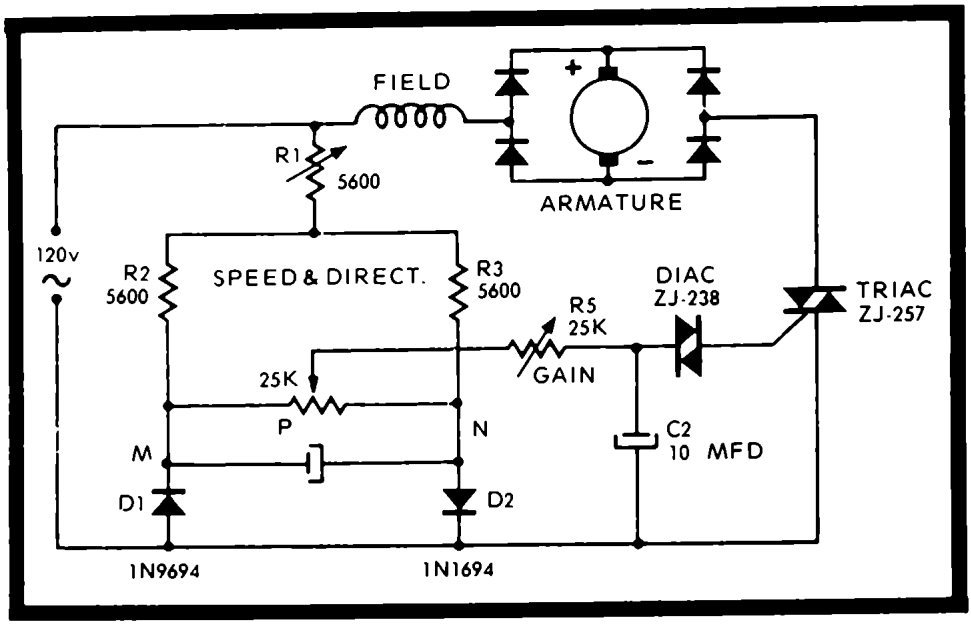


Fig. 8-56. Series motor reversing speed control circuit using a single Triac.

both SCRs. By means of potentiometer P, the quiescent bias of the UJT is set more or less below the peak-point firing voltage. With R1 equal to R2, the bridge will be balanced, the UJT will not fire and no voltage appears across the armature, so the motor stands still. Increasing R1 or decreasing R2 unbalances the bridge, an AC signal appears at the emitter of the UJT, biased above the firing voltage during one AC line half-cycle. The UJT fires and triggers SCR2, which happens to be forward biased. Decreasing R1 or increasing R2 would fire SCR2 instead of SCR1, reversing the polarity of the voltage across the armature, hence the rotating direction. If R1 and-or R2 is a suitable resistive transducer such as a photoconductor, a thermistor or a position-transmitting potentiometer, the motor speed and direction can be controlled by the physical quantity to be regulated.

Replacing two back-to-back connected thyristors, the Triac (Chapter 3) offers interesting shortcuts in circuitry. The circuit in Fig. 8-56 will reverse a series motor speed control with a single Triac. The main AC circuit comprises the motor field, a rectifier bridge with the armature connected across the DC diagonal, and the Triac. The latter is fired through a Diac by a signal sampled on potentiometer P, connected across diagonal MN of a rectifier bridge D1, D2, R2, and R3. In this bridge diodes D1 and D2 return to the lower bus with opposite polarities, so that positive half cycles appear on one end of P and negative ones on the other end. At the electrical midpoint of P, there is no signal, and the Triac does not con-

duct, thus failing to energize the motor. Leaving this midpoint, the slider samples a voltage that increases toward both ends and which is composed of half cycles of positive or negative polarity according to the direction of slider displacement. This increases the conduction angle of the Triac for positive or negative half cycles. While the armature is always fed with DC of unvarying polarity, the field will pass positive or negative half cycles of increasing amplitude, causing the motor to rotate in one direction or the other with increasing speed. Resistors R1 and R5 allow adjustment of the control "gain."

The split-field motor used for control purposes is a series motor having a center-tapped field winding, with the midpoint tied to one commutator brush. Such a motor can be reversed by a single-pole double-throw (SPDT) switch (Fig. 8-57), and it is easily fitted with a reversing electronic speed control using two thyristors, SCR1 and SCR2, triggered by a single unijunction transistor through transformer Tr (Fig. 8-58). Regarding power supply transformer Ta, the thyristor polarities are in opposition. SCR1 can be fired only if Terminal X of the secondary of Ta is positive regarding Point Q; if it is negative, only SCR2 can be made to strike. The UJT is synchronized with the power line frequency to generate one pulse per cycle that will trigger only SCR1 or SCR2 according to its occurrence during the first or the second half cycle. If SCR1 conducts, the motor will be energized by Terminals X and Y and rotate in a given direction; triggering SCR2 instead

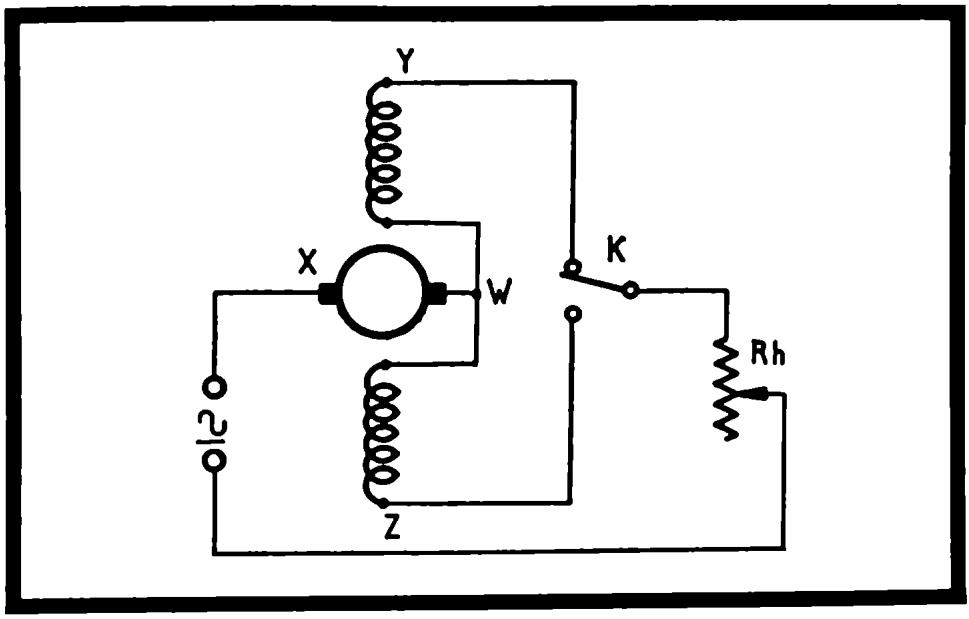


Fig. 8-57. A split-field motor can be reversed by a simple SPDT switch.

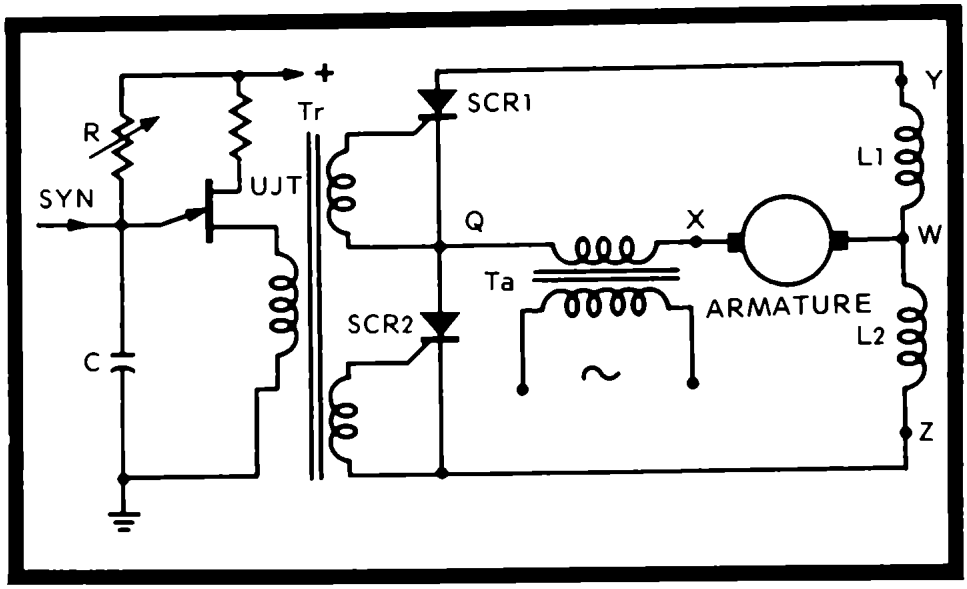


Fig. 8-58. UJT-triggered thyristor reversing speed control circuit.

during the remaining half cycles will supply the motor through Terminals X ans Z, causing it to rotate in the opposite direction. The motor speed can be varied by horizontal control.

A reversible split-field DC motor speed control using self-saturating magnetic amplifiers is schematically represented in Fig. 8-59. The control current traversing control windings Lc and L'c increases the apparent inductance of load windings L1 and L2 while decreasing the inductance of L'1 and L'2 (or vice versa), differentially varying the current in both halves of the split field. This will rotate the armature in one direction or the other, with the pace set by the control current. Such circuits are used in positioning devices.

In this type of circuit, both halves of the winding may pass equal though opposing currents, thus stalling the motor since the resulting field is zero. Such a situation will not be tolerated by conventional motors because of the low-resistance opposition to the applied voltage. A stalled motor can generate no cemf; therefore, excessive heat will develop, eventually leading to burn-out. The split-field universal motor, however, is designed to resist such punishment to allow for differential control.

ACCURATE DC MOTOR SPEED CONTROL

Low-power instrumentation types excepted, motors for industrial drives are of the shunt or separate-field sort which feature a convenient speed control characteristic. Fair speed regulation has been obtained in the foregoing, using the ar-

mature-generated counter emf to control the thyristor firing angle (Fig. 8-53). The accuracy of this simple approach, however, is limited by the fact that, in addition to the cemf, a voltage drop, Ri, proportional to current I traversing armature resistance R, develops across the armature. Because of Ri, this voltage is no longer proportional to motor speed.

The separate-field motor speed control in Fig. 8-60 features Ri-compensation. For this purpose, the armature current, rectified by full-wave connected thyristors SCR1 and SCR2, traverses the primary of transformer T1; this is an undulating current with an amplitude approximately proportional to I. The T1 secondary voltage is rectified and filtered. One output conductor is connected to the slider of "speed reference" potentiometer P2, providing a variable bias voltage, while the other is connected to the noninverting input of a differential amplifier in which the inverting input is connected to the slider of potentiometer P1 shunted across the armature and primary of T1. Since the voltages applied to these two inputs substract, the Ri component can be eliminated by a suitable setting of P1.

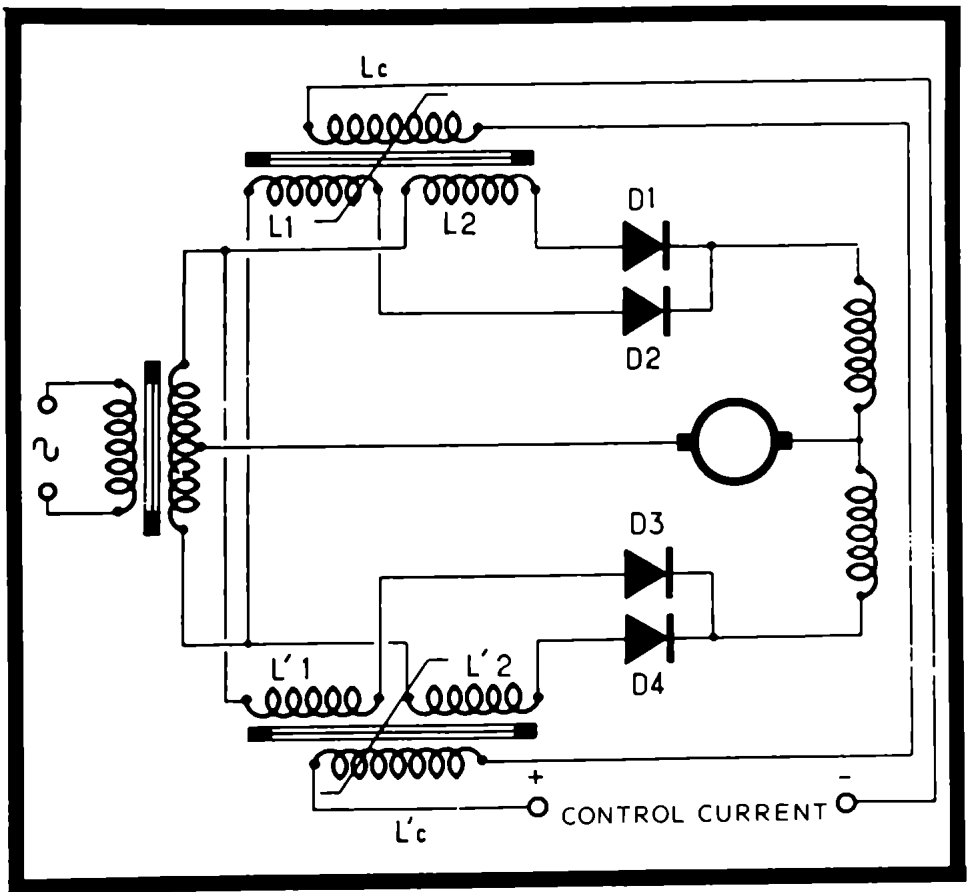


Fig. 8-59. Split-field motor reversing speed control circuit using a magnetic amplifier.

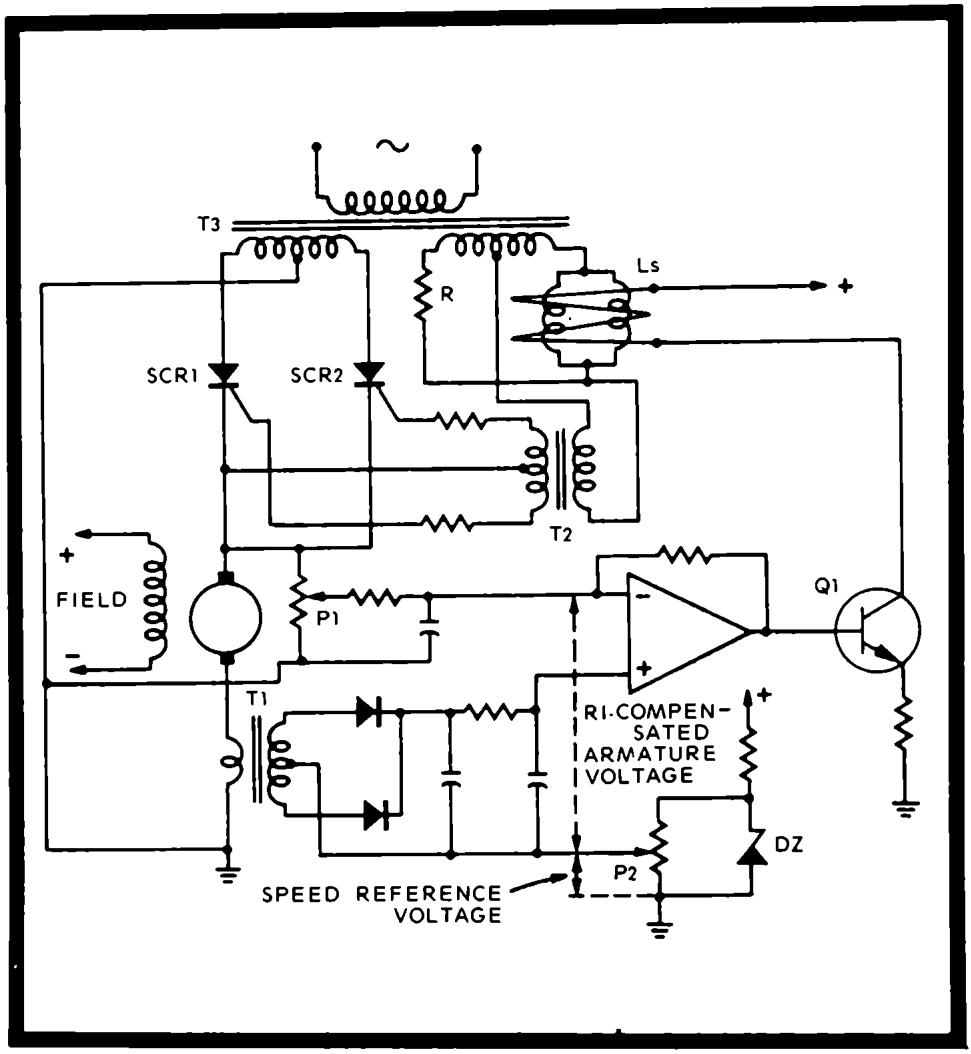


Fig. 8-60. RI-compensated precision speed control circuit for separate-field motors.

Transistor Q1, controlled by the differential amplifier, excites the control winding of saturable reactor Ls, which forms with resistor R a phasing bridge that sets the thyristor conduction angle. Thus, the speed programmed by setting P2 is maintained accurately, in spite of load and line voltage variations.

The full-wave shunt motor speed regulator in Fig. 8-61 features accurate wide-range speed regulation due to the control loop gain obtained by making use of the cemf and the inductive motor characteristics to control high-gain thyristor SCR. While the speed is continuously variable from a standstill to the top limit, regulation is effective for a speed ratio of at least 6:1, and it can be extended to cover a ratio of 20:1; that is, from 75 to 1,500 RPM, for instance. Power supply and load variations have no effect on the operating speed.

Zener diode DZ provides a regulated 25v supply for the UJT trigger circuit. At the beginning of the cycle, capacitor C charges through speed control R3 and diode D1 prevents the load current from diverting through thyristor SCR. When the UJT fires, a trigger pulse is transmitted by T1 to the SCR gate, which strikes the SCR and allows a current to pass through the armature. Since the field is permanently excited the motor starts rotating. The armature now develops a cemf and forward-biases D1 to divert the load current of C through R4. The amount of current that C is deprived of as a result depends on the cemf developed by the armature, hence on the motor speed, and also on the magnitude of R4 in relation to the motor characteristics.

Upon a torque increase, the motor has a tendency to slow down and drain more current. The slowed-down armature develops a smaller cemf, and the increased current drain increases the duration of the first part of the half cycle where the armature circuit closed through rectifier D2. These two effects both act to accelerate the charge on C, thus advancing the firing point of the SCR to pass armature current during a greater part of the half-cycle. More power is supplied to the armature, allowing it to meet the higher torque demand.

Very accurate speed regulation is obtained by using a tachometer generator; this is an instrument-type DC

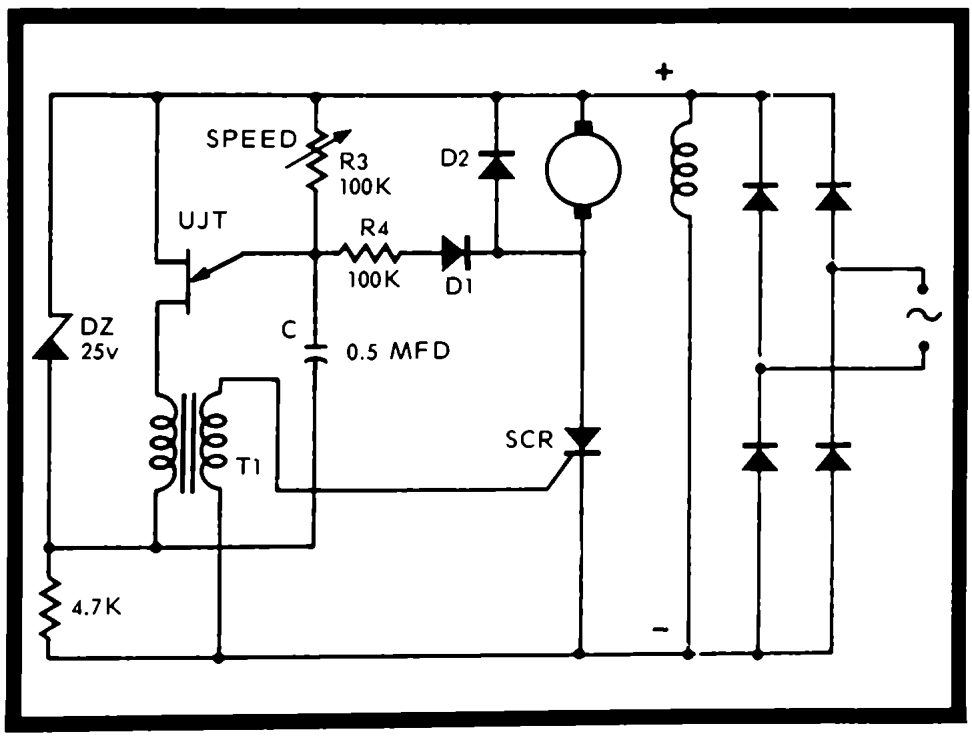


Fig. 8-61. Schematic of a wide-range, full-wave shunt motor speed regulator.

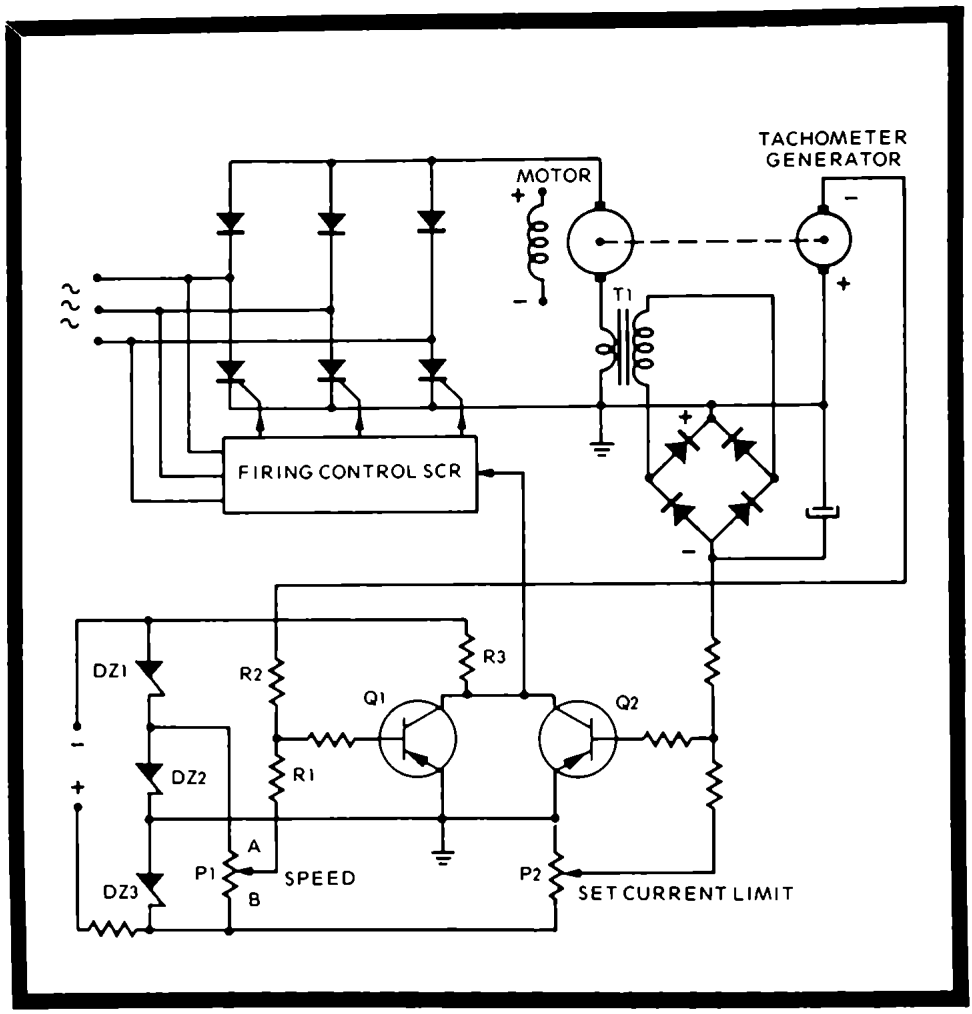


Fig. 8-62. Circuit using a tachometer generator for accurate speed control of large industrial motors.

generator designed to produce an accurate speed-proportional voltage. The tach is mounted directly on the shaft of the motor to be monitored. A tachometer generator-regulated separate-field motor for large industrial drives is schematically represented in Fig. 8-62. The 3-phase power is rectified by a bridge containing three conventional and three controlled rectifiers, horizontally controlled by an SCR control such as discussed in the foregoing to supply to the armature a continuously variable current between about zero and full power. (The constant-voltage DC field supply is not shown.)

The negative commutator brush of the tachometer generator is connected through resistor R2 to the base of transistor Q1. Q1's collector is connected through R3 to the negative busbar, which is tied to SCR control. Through R1, the base of Q1 is also connected to the slider of speed-setting potentiometer P1, which is connected across voltage-

regulating zener diodes DZ2 and DZ3 so that the slider can be above, below, and even at ground potential. At the lowest speed (slider at A), the base of Q1 is slightly negative, and its collector is, too. A very low voltage is applied to the SCR control (the thyristors do not supply armature current).

Rotating the slider of P1 toward B makes the base less negative, while the collector runs more negative. A higher voltage is applied to the SCR control, and the thyristors power the armature, which now starts rotating. The generator develops a speed-proportional voltage which is also applied to the base of Q1. That transistor compares the voltage with the preset speed signal and accordingly controls the armature current. If the motor begins to slow down because of a torque

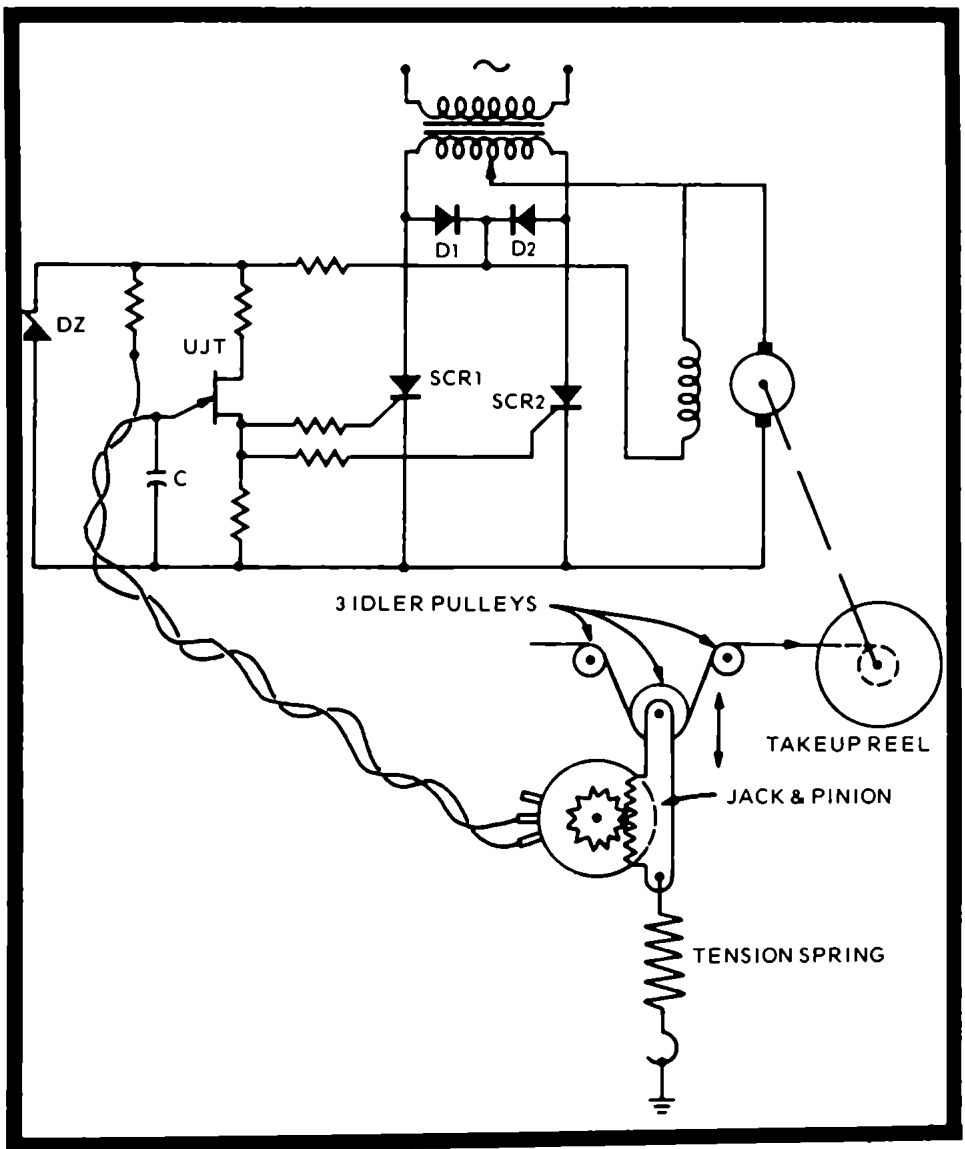


Fig. 8-63. Circuit designed to control motor speed by a production process variable.

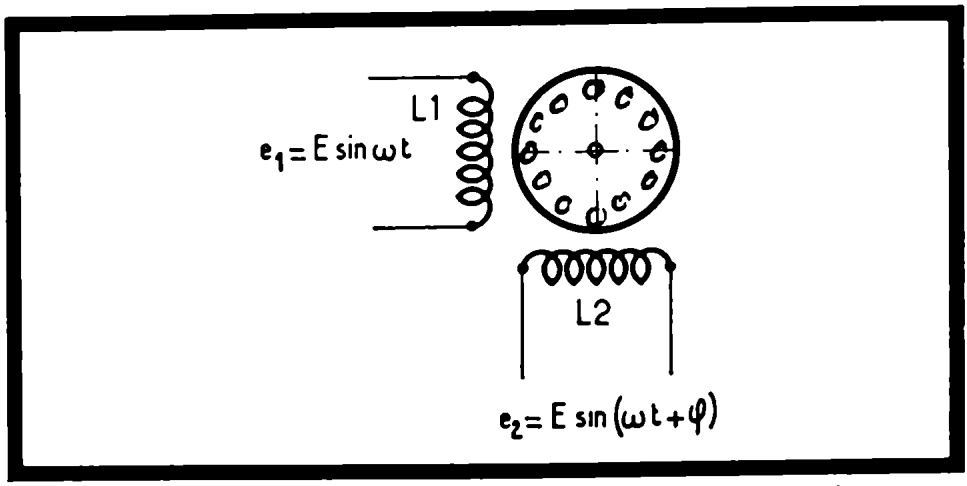


Fig. 8-64. A 2-phase induction motor contains two orthogonal windings and a squirrel-cage (brushless) rotor.

increase, the base becomes more positive, and the armature is fed an increased current to meet the load increase; this holds the motor speed at the preset level.

Transistor Q2 protects the motor against overcurrent. Though Q2 parallels Q1, it could pass unnoticed up to now because its positively biased base makes it normally nonconducting; this bias is provided by current-limit setting potentiometer P2. The armature current, however, traverses the primary of transformer T1 whose secondary supplies a rectifier bridge. This is a sort of DC transformer, for the primary is traversed by the undulating unfiltered rectified current. The smoothed output of the rectifier bridge is also supplied to the base of Q2. If the preset voltage provided by P2 fails to exceed this overcurrent signal, Q2 begins conducting and thereby increases the voltage drop across the common collector resistor R3; this reduces the SCR control voltage and the armature is fed less current.

Sometimes the operating speed of a driving motor must be controlled according to a process variable. For instance, in a wire drawing and varnishing plant, the tension of the processed wire must be kept constant. The takeup reel, therefore, gradually slows down as the diameter of the spooled material increases. According to Fig. 8-63, this tension can be sensed by a device composed of three idler pulleys; the two outer ones are fixed, while the center pulley is movable vertically, pulled down by a tension spring. Moving back and forth, the pulley drives a rack, thereby rotating a pinion connected to the axle of potentiometer P. A variation of wire tension thus results in a resistance variation that controls the charging current of capacitor C in the UJT emitter circuit and the trigger pulses by the same token. (For a discussion of this

circuit, refer to Fig. 8-20 and the related description.) The full-wave rectifier, D1 and D2, also provides the field voltage supply. Similarly, the motor speed can be made to depend on other parameters.

SQUIRREL-CAGE TWO-PHASE SERVOMOTOR CONTROL

The 2-phase induction motor used in servo controls consists of a stator with two sets of orthogonal pole windings, L1 and L2, and a laminated rotor stack with slots containing solid copper or aluminum bars joined at both ends by copper rings forming a sort of squirrel cage. If voltages presenting a quadrature phase relationship are applied to L1 and L2, a rotating field is set up by the stator, which induces voltages in the squirrel-cage conductors (Fig. 8-64). Since these conductors are shorted, they, in turn, produce a magnetic field that combines with the former to cause the rotor to rotate in the direction of the rotating field. This is an **asynchronous** motor, for its speed must remain below that of the rotating field; if that speed were reached, there would no longer be an induced voltage and, hence, a driving torque.

The driving torque of the 2-phase motor is proportional to voltages e1 and e2 across windings L1 and L2, and to the sine of the phase offset angle ϕ between e1 and e2. For quadrature phase offset, sin ϕ equals 1, and maximum torque is obtained. To reverse the motor, it is necessary only to commutate the phases or to reverse one winding.

To operate a 2-phase motor on a single-phase line, it is merely necessary to connect in series with one winding a capacitor (C) of adequate capacitance to shift the phase of the current in that winding by 90 degrees (Fig. 8-65). **Split-phase**

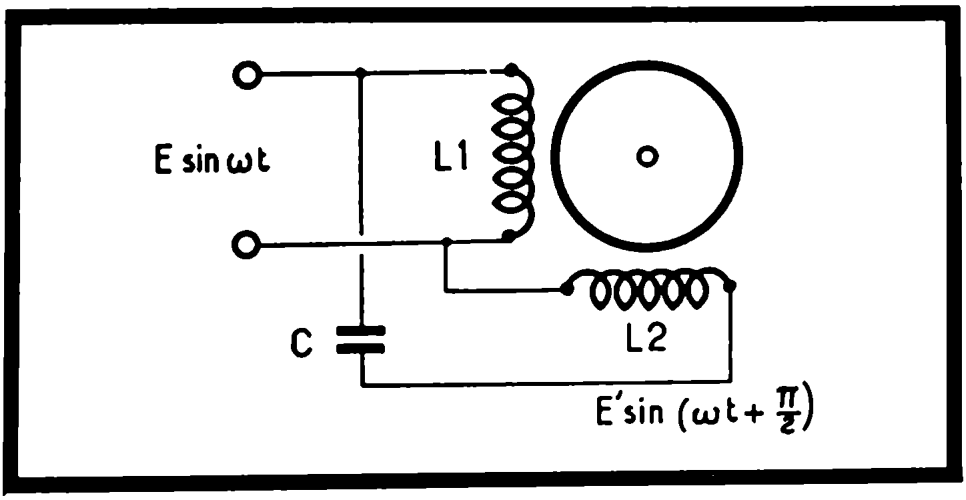


Fig. 8-65. For single-phase operation, a split-phase induction motor uses a phasing capacitor.

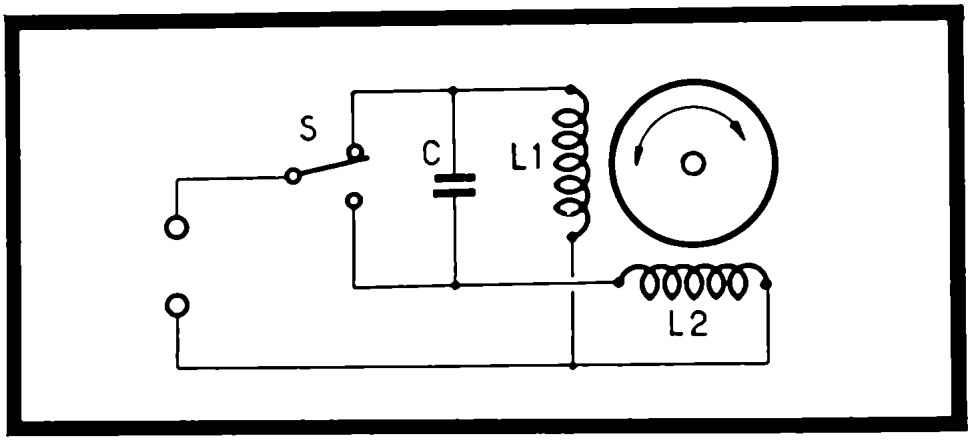


Fig. 8-66. Circuit showing how a split-phase motor can be reversed by an SPDT switch or relay.

induction motors are extensively used for instrumentation and control, for they provide application flexibility and match electronic amplifiers. If similar windings are used, the motor can be reversed by capacitor switching by means of a SPDT switch or relay (Fig. 8-66).

The **shaded-pole** motor is just another type of split-phase induction motor. Because of its built-in quadrature winding made up of one shorted turn of heavy copper ribbon encircling part of the pole pieces (Fig. 8-67), the direction of the rotating field is set once for all and the motor is not reversible. Two such motors, however, can be mounted on the same shaft and excited alternately or differentially to obtain rotation in both directions.

In squirrel-cage motors for instrumentation and control purposes, the two windings generally are designed to be

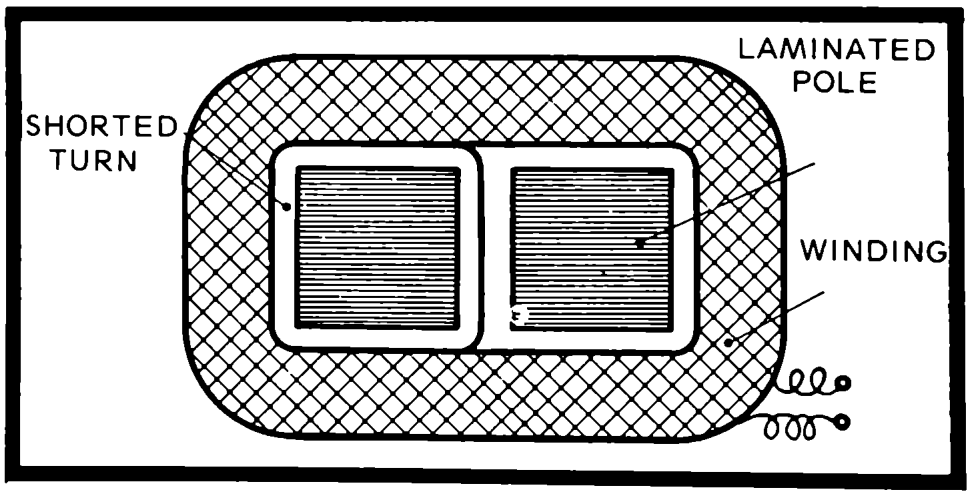


Fig. 8-67. Cross-section of the field coil in a shaded-pole induction motor. Such motors are not reversible.

dissimilar. The **control winding**, L1, is designed for low power requirements to ease the task of the instrument amplifier; the driving torque is mainly provided by the **reference winding**, L2, connected to the line or to the reference oscillator. The driving torque can be proportional to the error voltage phase and-or amplitude. The motor is designed to stand full voltages applied to both windings while stalling. The error voltage causes rotation in the appropriate direction; if the error voltage vanishes, or if the phase offset between e1 and e2 becomes zero (sin ϕ equals 0), the motor stops.

A bridge and a potentiometer, both self-balancing by means of split-phase motors, are discussed in Chapter 4 (Figs. 4-11 and 4-12) to illustrate the measurement principles. Fig. 8-68 represents a Wheatstone bridge composed of motor-controlled variable resistance R1, resistive transducer R2 and fixed (ratio) resistors R3 and R4. This bridge is fed by a low-frequency AC voltage (400 or 50-60 Hz). The "detector" diagonal is connected to the inputs of a differential amplifier whose symmetrical output energizes control winding L1. The reference winding is connected to the bridge frequency supply through phasing capacitor C, providing a 90-degree phase

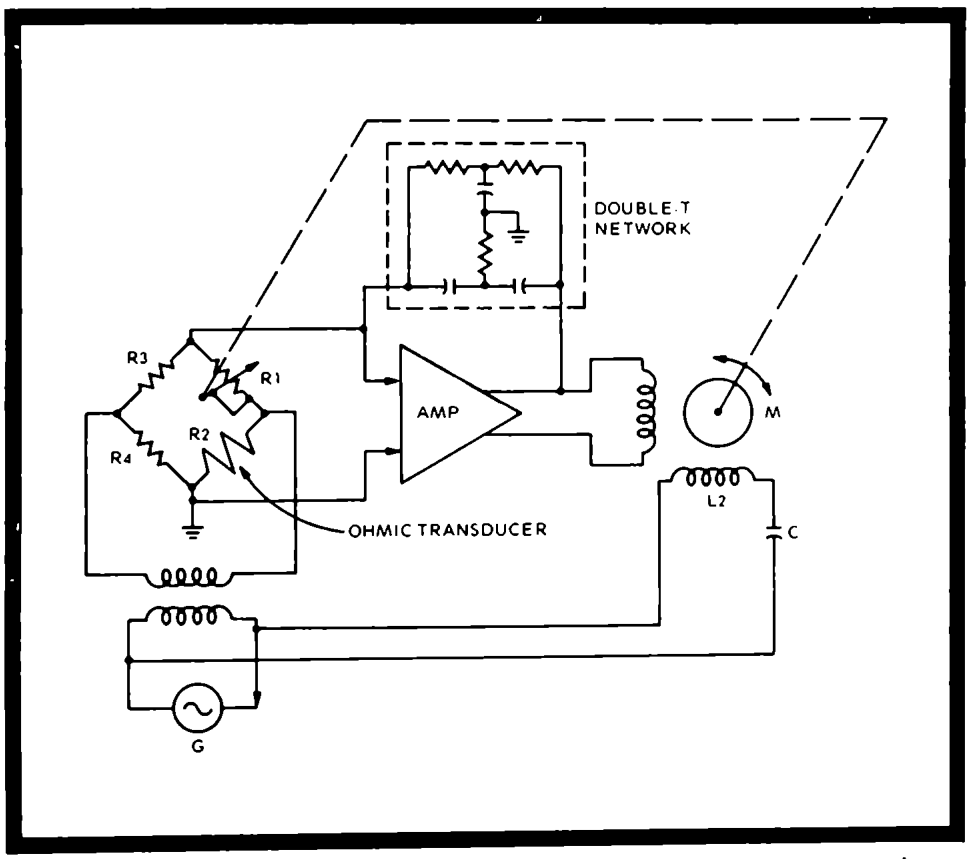


Fig. 8-68. Self-balancing bridge circuit using a 2-phase induction motor.

offset. Since the Wheatstone bridge is purely resistive, there is no phase shift between generator G and L1; but since the phase of L2 is offset by 90 degrees in regard to G, the field set up by L1 will lead or lag with respect to the field of L2 according to the actual direction of bridge unbalance, causing the motor to rotate in the direction appropriate to drive the slider of R1 in a way to null the error voltage. Once the bridge is balanced, the motor stands still.

The same principle, of course, applies to impedance bridges. However, the operating peculiarities stem from the phase sensitivity of the 2-phase motor. This circuit is fully responsive to the resistive (in-phase) error voltage component, while insensitive to the quadrature reactive component. Omitting phase offset capacitor C will, on the contrary, enable the sole reactive component to control motor rotation. This opens up interesting opportunities for impedance-bridge self-balancing systems.

The 60-Hz line seems to be a simple and economic solution for bridge supply; the amplifier, however, will pick up the undesired hum of the mains frequency just as well as the error signals therefore, screening as well as hum-bucking circuits must be provided. To avoid bridge nulling errors, it may be easier to supply the bridge with a different frequency of 400 Hz, for instance, and to build selectivity into the amplifier by connecting a double-T network tuned to the bridge frequency into the feedback loop. (Two-phase instrumentation motors are designed for a specific operating frequency; a 400-Hz model will perform poorly on 60 Hz.)

A squirrel-cage control amplifier (General Electric) is schematically represented in Fig. 8-69. It is designed for control of a 400-Hz motor with a center-tapped control winding for transformerless push-pull drive. The circuit requires 3w of exciting power. The control comprises an inverse-feedback stabilized input amplifier (Q1, Q2 and Q3), and an emitter-coupled symmetric driver stage Q4 and Q5 which drives common-base push-pull output stage Q6 and Q7 by means of transformer T1. This output stage is supplied with an unfiltered full-wave rectified 400-Hz voltage; thus, there will be a succession of positive half cycles (for NPN transistors) of 50 volts peak. The reference winding is directly connected to the 400-Hz source. Capacitors C1 and C2 tune the primary of T1 and L1, respectively, to 400 Hz.

Since the current gain of CB-connected transistors is less than one, voltage step-down transformer T1 provides some current gain. The output stage is Class AB-operated. In the large-signal condition, diode D1 becomes conductive, shorting the bias; amplifier operation thus proceeds towards Class B.

This amplifier is said to be inherently stable and with the specified silicon transistors is serviceable between -55 and +125 degrees C—provided the capacitors and other passive components will stand such a punishing environment.

WARD LEONARD DRIVES

In medium and large speed-regulated industrial drives such as used in rolling mills, the driving motor armature is directly connected to the armature of a DC generator driven by a constant-speed 3-phase AC motor. The electronic speed control applies to the generator field, which allows for very flexible speed regulation by wide-range motor armature voltage control. Furthermore, excess electric energy is fed back by the AC motor into the line. Such recovery is welcome not only because it cuts down the electricity bill, but also because it avoids costly dissipation of considerable losses transformed into heat.

In the Ward-Leonard system, a line-fed 3-phase motor (M1) is directly coupled to a DC generator G in which a

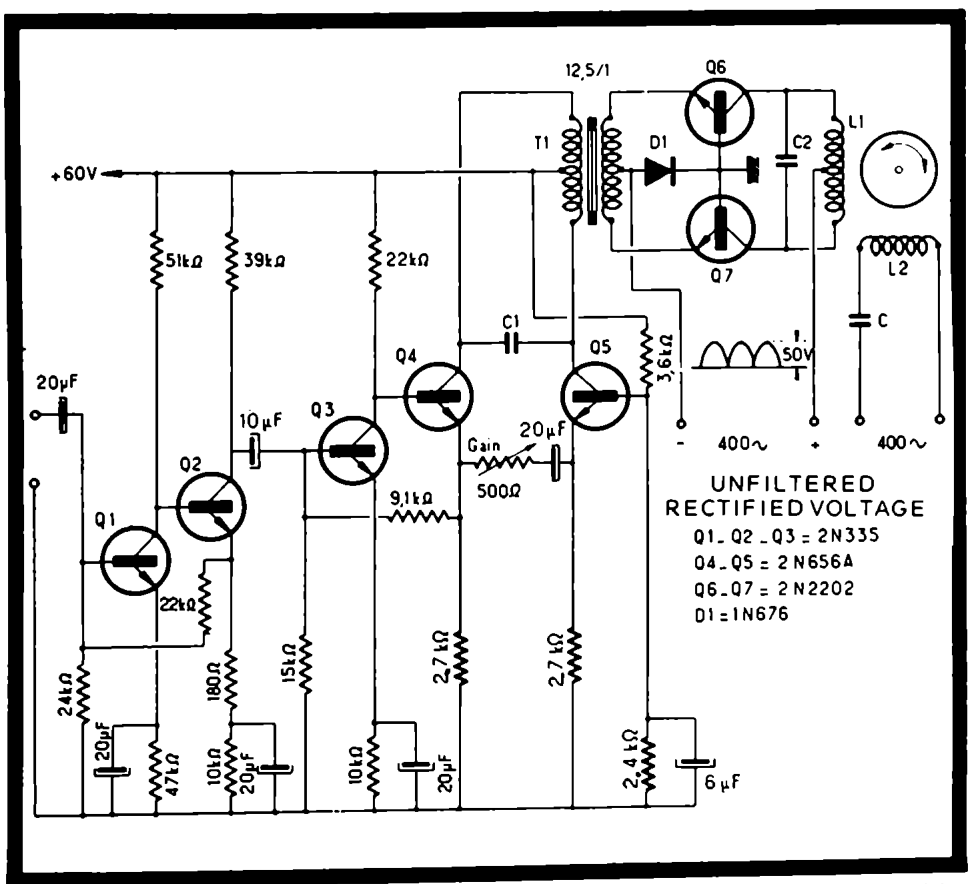


Fig. 8-69. Control amplifier circuit designed for a 400-Hz squirrel-cage motor.

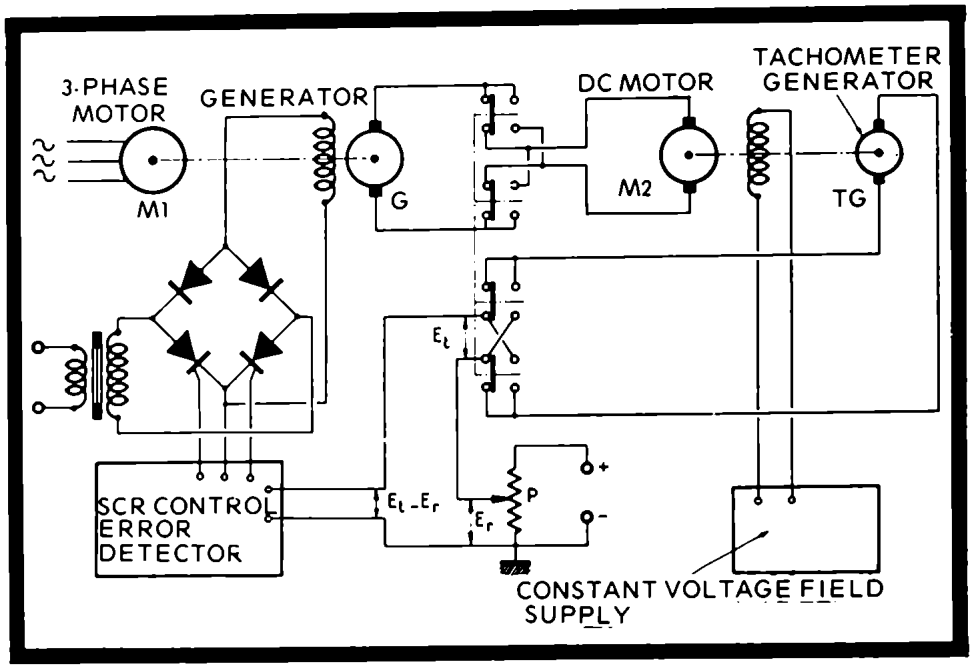


Fig. 8-70. Schematic of the Ward-Leonard system used for speed control of large industrial motors.

separate field is excited by a controlled rectifier bridge (Fig. 8-70). In large power systems, a pilot exciter generator is assembled on the same shaft and provides generator field voltage; the electronic regulation then applies to the exciter field. This motor-generator group can be set up in a generator plant at a distance from the mill-driving motor, M2, whose field is excited by a constant-voltage DC supply. Because of the direct connection of motor and generator armatures, the generator armature voltage (controlled by the generator field voltage) directly controls the motor speed. Since a comparatively small generator field voltage variation causes a considerable motor armature voltage variation, this, in effect, is a rotary amplifier.

A tachometer generator (TG) mounted on the shaft of driving motor (M2) provides a speed-proportional voltage (Et) opposed to a reference voltage (Er) that is adjustable by means of potentiometer P. The resulting error voltage (Et - Er) is applied to an amplifier which controls the conduction angle of the thyristor generator-field rectifier; thus the speed preset by P is maintained. If the drive is to be reversing, the M2 armature voltage is reversed by a contactor rather than its field voltage; in large shunt and separate-field DC machines, field polarity reversal is to be avoided. The same contactor simultaneously also reverses the tachometer generator output in order to maintain opposed Et and Er in spite of the reversed direction of rotation.

AMPLIDYNE ROTARY AMPLIFIER

The amplidyne is another type of rotary amplifier particularly suitable for large industrial drives. This is a special type of DC generator using, besides the normal set of brushes (A, B) that deliver the output voltage, an additional set of brushes (C, D) which are tied together (Fig. 8-71). Because of this short circuit, a large current is induced in the rotating armature. The resulting heavy current sets up a large cross flux perpendicular to the field flux. This cross flux induces in the rotating armature an output voltage available at brushes A and B to power the armature of the DC driving motor to be regulated. The generated voltage, of course, depends on the field excitation as in a conventional DC generator. However, because of the internally generated cross flux, a rather low field current variation of typically from 30 to 100 ma is sufficient to increase the armature output from 10 to 200 ohms. (A conventional DC generator would require fields roughly one hundred times stronger to achieve the same result.) While a gain (of field input-armature output power) from 25 to 100 is obtained with a DC generator, amplidynes feature gains of up to 10,000 and even more.

The load current produces an armature reaction flux that opposes the control field flux, thereby preventing control action. This effect is canceled by a winding located around the poles and connected in series with the armature to pass the full load current. Because of its low power requirements, the control field fairly well matches with electronic amplifiers. Furthermore, since these low-power control windings occupy

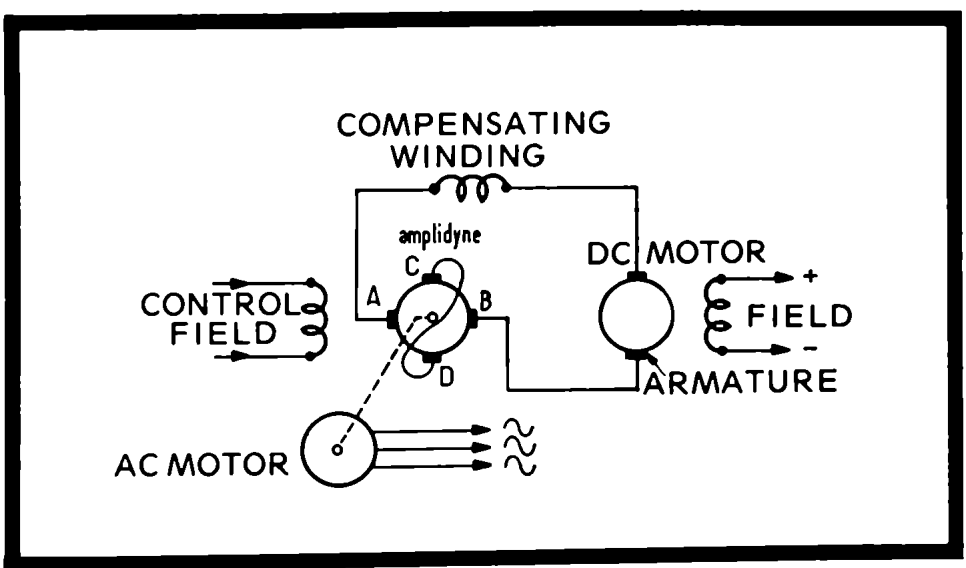


Fig. 8-71. Diagram of the connections in an amplidyne rotary amplifier.

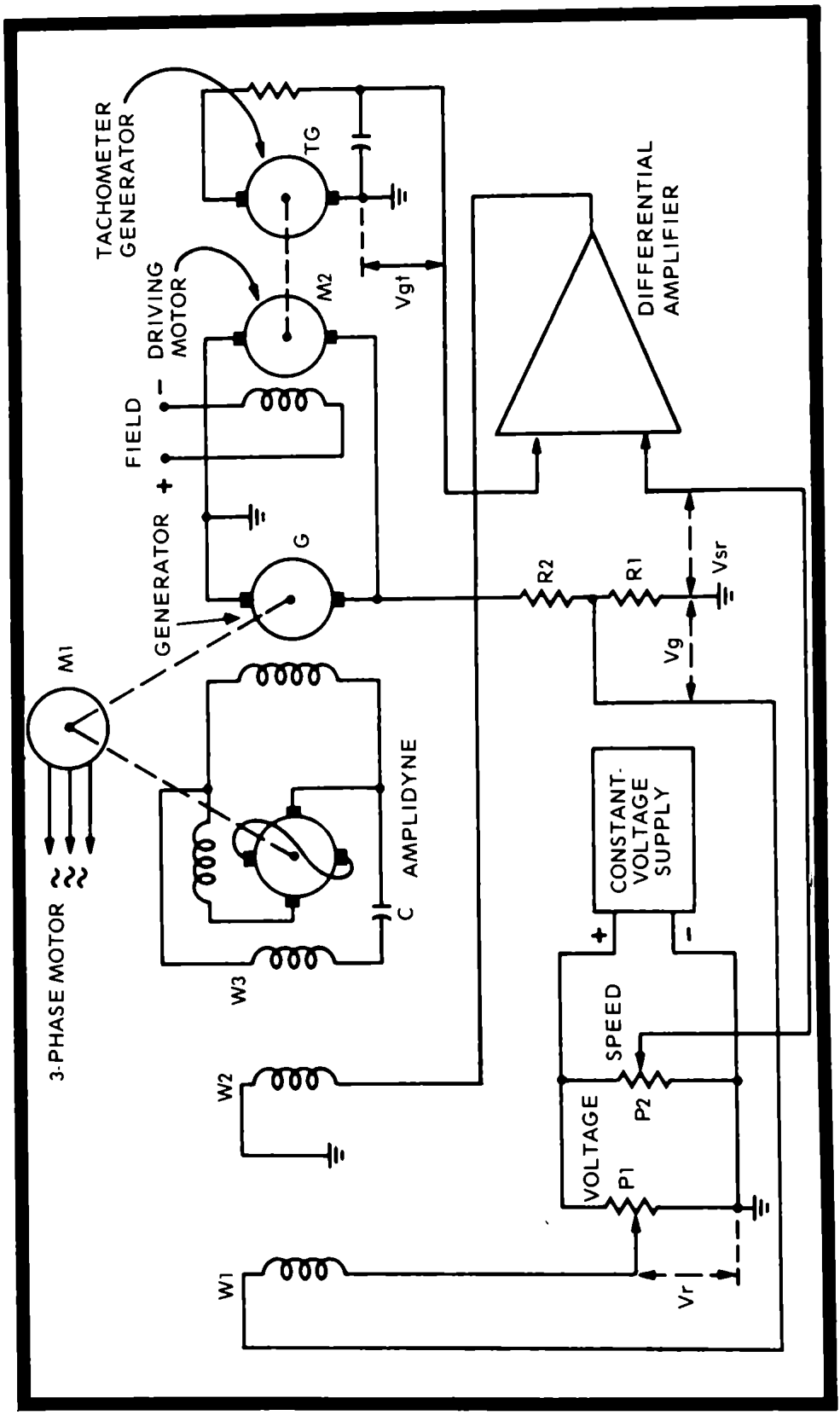


Fig. 8-72. Schematic of an amplidyne speed regulator for large industrial motors.

little space, several separate windings (generally 2 to 6) can be placed on each pole, allowing for refined control systems (somewhat similar to the magnetic amplifier technique).

An amplidyne-controlled speed-regulated large power industrial drive is schematically presented in Fig. 8-72. A 3-phase line-fed AC motor (M1) drives the amplidyne and a DC generator G, whose field is energized by the amplidyne output. The generator armature is directly connected to the armature of driving motor M2, which has a separate field supply. M2 coupled to a tachometer generator TG.

The amplidyne contains three control windings W1, W2 and W3. W1 regulates the generator voltage supplied to armature of motor M2. To this end, a generator-voltage proportional voltage, Vg, obtained from voltage divider R1-R2 is opposed to reference voltage Vr set by constant-voltage fed "voltage" potentiometer P1; the resulting error current traverses W1. When Vg equals Vr, the winding passes no current; however, only a slight unbalance is necessary to reinforce or to reduce the existing control field in one direction or the other.

The tachometer generator coupled with the armature of motor M2 generates a speed-proportional voltage, Vgt, which is applied to one input of a differential amplifier. The second input is a speed reference voltage, Vsr, from "speed" potentiometer P2, which is connected across the same constant-voltage supply as P1. The amplifier output controls W2, the main control winding of the amplidyne system. Therefore, the circuit has to be designed so that W2 always passes a current of unchanging polarity to avoid accidentally reversing the polarity of the motor field and the direction of motor rotation by the same token. This can be achieved by a superimposed direct current, for instance. If the motor slows down, Vgt decreases, W2 passes a higher current, the amplidyne provides more current to the generator field, and the generator feeds more current to the motor armature to accelerate it.

Contrary to a DC generator, an amplidyne immediately responds to any control current variation, possibly leading to unstable operation, the so-called "hunting" effect. W3 is not a control but an **anti-hunt** winding. Capacitor C transmits to W3 the output voltage transients caused by rapid control field variations. The field of W3 thus opposes these variations and introduces sufficient damping to avoid hunting. C, of course, transmits only the transients, not the DC component.

Index

A

AC amplifier, 128
Acidity measurements, 54
AC tachometer generator, 75
Adaptive controls, 278
Adder servomechanism, 284
Alkalinity measurements, 54
Ambient conditions, 10
Amplidyne rotary amplifier, 409
Amplifiers, vacuum tube, 108
Analog computers, 280
Analog phase inverter, 282
Analog-to- digital conversion, 270
AND function, 236
AND gate, diode, 249
AND gate, transistor, 245
Angular position synchro, 68
Automatic adder servomechanism, 284

B

Binary-coded decimal counter, 228
Binary counter, 257
Binary counting, 223
Blocking oscillator, 153
Boolean algebra, 236
Bootstrap circuit feedback, IC, 128
Bridged-T network oscillator, 147
Bridge modulator, 134
Bridge motor control, 406
Bridge rectifier, 346
Bridges, 33

C

Cable testing, 190
Capacitive chopper, 132
Capacitive transducers, 56
Capacitor testing, 196
Carrier-current amplifier, 129
Cavitation, 315
CE amplifier transistor, 112
Chopper modulator, 131
Cleaning, ultrasonic, 315
Clock generator, 218
Coefficient potentiometer, 281
Cold cathode thyratrons, 164
Collector, transistor, 110
Common collector, 111
Common emitter, 111
Complementary, transistor, 112
Conductivity, liquid, 53
Constant-current power supply, 362
Constant-load regulator, 382
Contactor, welding control, 307
Continuous and discontinuous controls, 13
Contouring system controls, 278
Controlled rectifiers, 350
Converters, 366
Coordinate positioning, 276
Counter, binary, 257
Crystal oscillator, 148
Current amplifier, 115
Cyclic code, 272

D

Dark current, photoemissive tube, 77
d'Arsonval meter, 33
DC motor control, 389
DC tachometer operator, 73
Dekatron, 212
Diac, 175
Dielectric heating, 296
Dielectric losses, heating, 296
Differential input null detector, 128
Differential receiver, 70
Differential transmitter 70
Differentiation circuits, 156
Differentiation, 160

Digital integrated circuits
Digital voltmeters, 219
Dimensional inspection, 203
Diode AND gate, 249
Diode logic (DL), 248
Diode matrix, 262
Diode-transistor logic (DTL), 248
Direct-coupled tube amplifiers, 109
Direct-coupled transistor, 109
Displacement transducer, 103
Doubler, voltage, 348
Double-T network oscillator, 147
Drilling, laser, 322
Dynamic strain gauge measurements, 49

E

Eccles-Jordan counter stages, 226
Eccles-Jordan multivibrator, 151
Electroluminescent diodes, 86
Electromechanical counters, 206
Electron-beam heating, 304
Electronic timer, 163
Electron tubes, 109
Evaporation, laser, 322
Events-per-unit-time counter, 218

F

Feedback, 26,107
Feedback, magnetic amplifier, 385
Feedback,oscillators, 143
Ferrite core matrix, 267
FETs, 116
Fiber optics, 86
Field-effect transistors, 116
Filter, power supply, 345
Fire detector, 332
Flame-failure protection, 330
Flip-flop, 257
Flip-flop oscillator, 150
Four-layer diode, 167
Free-running multivibrator, 149
Full-wave rectifier, 346

G

Gain, 105
Galvanometric amplifier, 200
Gamma radiation monitor, 339
Gas-filled rectifiers, 345
Geiger tube, 339
Gray code, 272

H

Half-wave magnetic amplifier, 383
Half-wave rectifiers, 344
Half-wave series motor control, 391
Hall-effect transducers, 101
Hall-effect voltage multiplier, 287
Heating, RF, 294
High-frequency heating, 294
Homing, 188
Hot cathode thyratrons, 162
Hybrid ICs, 122
Hysteresis losses, heating, 294

I

Ignitrons, 370
Impedance bridge, 35
Impedance converters, 141 , 282
Induction heating, 299
Inductive rotation transducer, 72
Inductive transducers, 63
Inductor testing, 196
In-line numeral displays, 212
Input capacitor, power supply filter, 345
Integrated circuits, 120
Integrating circuits,, 156
Interdiction logic, 242
Internal feedback, 113
Intruder alarms, 336
Inverse feedback, 26, 107
Inverters and converters, 366

J

Junction FET, 116
Junction photoelectric devices, 38

L

Laser beams, 322
LC oscillators, 148
Light-activated silicon controlled rectifier, 174
Linear ICs, 123
Line-voltage regulation, 362
Liquid conductivity, 53

Liquid processing, ultrasonic, 315
Liquid resistivity, 53
Logic diagrams, 241
Logic operations, 236

M

Machine control programming, 273
Machining, spark, 320
Machining, ultrasonic, 312, 317
Magnetic amplifier, 382
Magnetic core memory, 266
Magnetic hysteresis, heating, 294
Magnetic signal amplification, 266
Magnetic switching, 264
Magnetostrictive ultrasonic transducers, 314
Mean time between failures, 11
Memory, magnetic core, 266
Mercury-vapor rectifiers, 345
Microelectronics, 120
Monolithic ICs, 120
MOSFET, 119
Motor control, DC, 389
Motor driven, line-voltage regulator, 362
Motor speed, control, 173
Multivibrator, 148
Mutual inductance transducers, 66

N

NAND function, 240
Negative logic, 246
Nixie tubes, 213
NOR gate, transistor, 249
NOT function, 237
NOT gate, transistor, 249
NPN transistor, 112
Null indicators, 31
Numeral display devices, 210
Numerical position control, 276

O

Optical relay, 85
Opto-electronic chopper, 137
Opto-electronic modulator, 96
Opto-electronics, 85
OR function, 236
Overrating, 9
Overshoot, positioning control, 24

P

Peak inverse voltage, rectifier, 344
Pentode, amplifier, 109
Pentode tube, 109
Phanotrons, 345
Phase inverter analog, 282
Phase inverters, 157
Phasing bridge, 157
Phasing circuits, 156
PH measurements, 53
Photoconductors, 82
Photodiodes, 83
Photoelectric controls, 87
Photoelectric dimension sensor, 203
Photoelectric safety device, 326
Photoelectric transducers, 76
Photoemissive cells, 76
Photomultiplier tube, 79
Photothyristor, 92, 174
Phototransistor, 84, 87
Photovoltaic cells, 80
Piezoelectric ultrasonic transducers, 313
Power supply regulation, 359
PNP transistor, 112
Potentiometer voltage measurement, 36
Positioning control, 23
Position transducer, 39
Positive logic, 245
Positive protection safety devices, 328
Post-compensation, IC, 127
Power amplifier, 106
Precompensation, IC, 128
Preset binary counter, 259
Preset counter, 215
Processing, ultrasonic, 312
Production control, 8
Production line counting, 231
Program control, 273
Programmed control, 22
Proximity switch, 337
Pulsed laser beams, 325

Punched card programming, 273
Punched tape program control, 274
Punch press safety device, 326

Q

Quality control, 205

R

Radiation monitor, 339
Raysistor, 96
RC oscillators, 148
Rectifiers, 343
Redundancy, 11
Regulated power supplies, 359
Regulator, electronic, 20
Regulator, speed governor, 20
Relaxation oscillators, 148, 177
Reliability, 9
Resistance welding controls, 306
Resistive tranducers, 39
Resistive temperature transducer, 50
Resistor-capacitor-transistor logic (RCTL), 248
Resistor testing, 192
Resistor-transistor logic (RTL), 248
Response, magnetic amplifier, 389
Reversible split-field motor, 396
Reversing DC motor control, 392
Revolution counting, 72
RF heating, 294
Ring modulator, 67
Rotary amplifier, 409

S

Saturable reactor regulator, 364
Saturable reactors, 379
Schmitt trigger, 151, 257
SCR, 170
Screen grid, 108
Selenium photovoltaic cells, 81
Self-balancing bridges and potentiometers, 197
Self-saturating magnetic amplifier, 385
Separate-field DC motor, 389
Series DC motors, 389
Servomechanism, 20
Shaded-pole motor, 404
Shear safety device, 326
Shunt DC motor, 389
Shunt motor speed regulator, 396
Signal-polarity inverter, 127
Silicon controlled rectifier, 170
Silicon photovoltaic cells, 81
Single-phase half-wave rectifier, 344
Single-shot multivibrator, 150
Smoke detector, 332
Soldering, ultrasonic, 317
Solid-state rectifiers, 343
Spark machining, 320
Speed control, 173
Split field DC motor, 395
Split-phase motor, 403
Spot welding, 307
Squirrel-cage motor control, 404
Squirrel cage 2-phase motor, 403
Stability, servomechanism, 26
Static strain measurements, 44
Static switching, 243
Stepping switch, 187
Strain gauges, 42
Suppressor, 109
Suppressor grid, 109
Switching characteristics, transistor, 139
Synchro comparator, 69
Synchros, angular position, 68
Synchronous welding control, 311

T

Tachometer generator, 73
Tachometer generator motor control, 408
Tachometer motor-speed regulator 397
Temperature compensation transistor, 116
Temperature measurement, 50
Tetrode tube, 109
Thermal compensation of direct-coupled transistor amplifiers, 116

Thermal [illegible] sistor, 116
Thermistors, 51, 181
Thickness gauge, 61
Three-phase controlled rectifier, 357
Three-phase full-wave bridge rectifier, 349
Three-phase rectifier, 349
Thyratrons, 162
Thyristor, 91, 167, 356
Thyristor motor control, 391
Thyrite, 181
Time constant, magnetic amplifier, 389
Timer, welding, 309
Torque measurements, 49
Transient behavior, 23
Transistor amplifier, 112
Transistor AND gate, 245
Transistor chopper, 134
Transistor NOR gate, 249
Transistor NOT gate, 249
Transistors, 109, 110
Transistor switching characteristics, 137
Transistor-transistor logic (TTL), 245
Triac, 174, 377
Triode amplifiers, 108
Trochotron, 211
Tube amplifiers, 108
Tube rectifiers, 344
Tuning fork oscillator, 148
Two-phase induction motor, 403

U

Ultrasonic cleaning and liquid processing, 315
Ultrasonic machining, 317
Ultrasonic processing and machining, 312
Underrating, 10
Unijunction transistors, 176
Univibrator, 150

V

Vacuum photocells, 79
Vacuum tube, 108
Vacuum tube amplifiers, 108
Variable reluctance transducers, 63
Varistor, 181
Vibrating-reed chopper, 131
Voltage amplifiers, 105 , 115
Voltage doublers, 348
Voltage feedback, 107

W

Ward-Leonard drive, 407
Weight, binary digits, 229
Welding, laser, 322
Welding timer, 309
Welding, ultrasonic, 317
Wheatstone bridge, 33
Wheatstone bridge motor control, 405
Wien bridge oscillator, 148

X

X-ray monitor, 339

Z

Zener diode, 182